

Db 7103 TGCTGCTCCGAAGCCATAGGAAGTTGCAAAATCCCGGCACTGAGAAAGAAAACAGCTTC 7162
Qy 61 TTGAATTTCTTGAAGTACCTCTTAATTCATTCATGCTGACCTCGGAGATGTAGTAG 120
Db 7163 TTGAATTTCTTGAAGTACCTCTTAATTCATTCATGCTGACCTCGGAGATGTAGTAG 7222
Qy 121 CTGACTCATTAATTCACAGGTGTTATGTTCTACCCATTCATCTGCAACATGAATTAATG 180
Db 7223 CTGACTCATTAATTCACAGGTGTTATGTTCTACCCATTCATCTGCAACATGAATTAATG 7282
Qy 181 ATTCTGAAACCACTGCCCCCTAGGGGAACTACAAGGTATGATTCGCTGAGCTCTGCTCA 240
Db 7283 ATTCTGAAACCACTGCCCCCTAGGGGAACTACAAGGTATGATTCGCTGAGCTCTGCTCA 7342
Qy 241 CGTTTGTGTTAACCAATCAATTAATTAATACCTTGTGTTGTGTCATTTCTGTTTAAATAT 300
Db 7343 CGTTTGTGTTAACCAATCAATTAATTAATACCTTGTGTTGTGTCATTTCTGTTTAAATAT 7402
Qy 301 CTTTATTTAATAGCACTGCTTAATTTCTTCAACATTTGTTTCAAAACCAAAAGGCCCTATT 360
Db 7403 CTTTATTTAATAGCACTGCTTAATTTCTTCAACATTTGTTTCAAAACCAAAAGGCCCTATT 7462
Qy 361 AACTGGAAGCCCTGAATGAAGCTTACATACACACATTTGTTTCTATAGAGAAAATTT 420
Db 7463 AACTGGAAGCCCTGAATGAAGCTTACATACACACATTTGTTTCTATAGAGAAAATTT 7522
Qy 421 TTCTTCAATCTCTGCCACAGCTTCTGCTTAAATGTGGAACAATATATACATACATG 480
Db 7523 TTCTTCAATCTCTGCCACAGCTTCTGCTTAAATGTGGAACAATATATACATACATG 7582
Qy 481 AAATGCACTATTTAATCTAATTTTAACTGTACAGTTCAGTGCATTAATACATACCTTAC 540
Db 7583 AAATGCACTATTTAATCTAATTTTAACTGTACAGTTCAGTGCATTAATACATACCTTAC 7642
Qy 541 CTGCTGTGAACTATGATCAACATTCATCTCTAGAACTTTTGTATCTTCTGCAATGAA 600
Db 7643 CTGCTGTGAACTATGATCAACATTCATCTCTAGAACTTTTGTATCTTCTGCAATGAA 7702
Qy 601 AACTGTGATCTTATGCAACATGAGCTTTCCCAATCACACCGACCTTCTTCCAGCCCC 660
Db 7703 AACTGTGATCTTATGCAACATGAGCTTTCCCAATCACACCGACCTTCTTCCAGCCCC 7762
Qy 661 CCGGACACACATCTTCTTCCATGACATGCTGCTGTGGAACAAGGCCCTTCA 720
Db 7763 CCGGACACACATCTTCTTCCATGACATGCTGCTGTGGAACAAGGCCCTTCA 7822
Qy 721 ACACTAGCTTGGGGGCACTGTAGCAACAGATCACTCCCTACCGGCAACAATGAC 780
Db 7823 ACACTAGCTTGGGGGCACTGTAGCAACAGATCACTCCCTACCGGCAACAATGAC 7882
Qy 781 ACAAAAATATTAAGCAATGCTGATTCGATTCGAAAAGGCTGCTTGTAAAGTATGA 840
Db 7883 ACAAAAATATTAAGCAATGCTGATTCGATTCGAAAAGGCTGCTTGTAAAGTATGA 7942
Qy 841 GGGGTGAAGCAAGGCAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 900
Db 7943 GGGGTGAAGCAAGGCAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8002
Qy 901 AGCTCAAAATTTTCCATTTGTTCTGTGCTAGCAACAATGCTTATTAAGCACTGTAAAG 960
Db 8003 AGCTCAAAATTTTCCATTTGTTCTGTGCTAGCAACAATGCTTATTAAGCACTGTAAAG 8062
Qy 961 ATTGATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Db 8063 ATTGATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8122
Qy 1021 GAATTAAGAGCACTGACATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1080
Db 8123 GAATTAAGAGCACTGACATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8182
Qy 1081 AGGAGAAAGGCTCAGAGGCTGCTTACACATTTTATCTAGAGAGCTAGGCTATGCTAG 1140
Db 8183 AGGAGAAAGGCTCAGAGGCTGCTTACACATTTTATCTAGAGAGCTAGGCTATGCTAG 8242

Qy 1141 TCGGTAGCTGGCAATTAACAACTGAAGCACTTCTAGTTCATGTGAGAGATGAATTTAAC 1200
Db 8243 TCGGTAGCTGGCAATTAACAACTGAAGCACTTCTAGTTCATGTGAGAGATGAATTTAAC 8302
Qy 1201 ATAACTCAACCCCTCTGCACTGAAGACAGACTAAGTCACTAAGTCACTAAGTCACTA 1260
Db 8303 ATAACTCAACCCCTCTGCACTGAAGACAGACTAAGTCACTAAGTCACTAAGTCACTA 8362
Qy 1261 ACTTACTATATGACAGGCACTGATCTGACAACTTACATGATTAATTAATTAATTAAT 1320
Db 8363 ACTTACTATATGACAGGCACTGATCTGACAACTTACATGATTAATTAATTAATTAAT 8422
Qy 1321 CCAACACTCTATGAGAACTGAAGTGAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGA 1380
Db 8423 CCAACACTCTATGAGAACTGAAGTGAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGA 8482
Qy 1381 CAAACGTAGATCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1440
Db 8483 CAAACGTAGATCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8542
Qy 1441 CAATTTCAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500
Db 8543 CAATTTCAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8602
Qy 1501 TATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
Db 8603 TATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8662
Qy 1561 TAAACCTTTAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1620
Db 8663 TAAACCTTTAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8722
Qy 1621 GGGCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1680
Db 8723 GGGCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 8782
Qy 1681 ATGTAAGGCCATTTTCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 8783 ATGTAAGGCCATTTTCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 8842
Qy 1741 TTAACTTTTCAAGGCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1800
Db 8843 TTAACTTTTCAAGGCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8902
Qy 1801 TGTCGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1860
Db 8903 TGTCGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8962
Qy 1861 GAAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1920
Db 8963 GAAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 9022
Qy 1921 AATAAATCTGCTGAGGCTGGAAGACGCTTCTGTTCTTGTGCTGATTAATTTGTTAGT 1980
Db 9023 AATAAATCTGCTGAGGCTGGAAGACGCTTCTGTTCTTGTGCTGATTAATTTGTTAGT 9082
Qy 1981 TTTTAAATAGTTCATTTATTAAGACCTTGTGAGCTCCCAAGCTAAGTGAAGTGAAGTGA 2040
Db 9083 TTTTAAATAGTTCATTTATTAAGACCTTGTGAGCTCCCAAGCTAAGTGAAGTGAAGTGA 9142
Qy 2041 CTTACAGAGGCACTTCAATTTTGAAGAACAAAAAGCCCAATTTCTGCTCCAGACCTTGA 2100
Db 9143 CTTACAGAGGCACTTCAATTTTGAAGAACAAAAAGCCCAATTTCTGCTCCAGACCTTGA 9202
Qy 2101 CCCCCAATCTCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 9203 CCCCCAATCTCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9262
Qy 2161 TAGAGGTGCGACAGTACTAGAAATGAAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 2220
Db 9263 TAGAGGTGCGACAGTACTAGAAATGAAGAAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 9322

QY 2221 AAAATGGGACCCCTTCATGATGCTCTTCCCTTCGTGTTCAAACTGCCATA 2280
 Db 9323 AAAATGGGACCCCTTCATGATGCTCTTCCCTTCGTGTTCAAACTGCCATA 9382
 QY 2281 TAATTTTACAGAGGTATATACATCTTTTCTTTTACACCTGGAAGGTGAGAAA 2340
 Db 9383 TAATTTTACAGAGGTATATACATCTTTTCTTTTACACCTGGAAGGTGAGAAA 9442
 QY 2341 CTTTGTTACCCATCATATAAATTCATCTTCTTAAGTCATCTTAATTTCTAAGATCA 2400
 Db 9443 CTTTGTTACCCATCATATAAATTCATCTTCTTAAGTCATCTTAATTTCTAAGATCA 9502
 QY 2401 AATAGCTGACAAATATCTCTTTTGTAAATTAACAATGAAAAACATCTCTGACAAATAT 2460
 Db 9503 AATAGCTGACAAATATCTCTTTTGTAAATTAACAATGAAAAACATCTCTGACAAATAT 9562
 QY 2461 TAATCTGCACTTTAGATATAGAAATTAATTAATCTAATGAAATGAAATCA 2520
 Db 9563 TAATCTGCACTTTAGATATAGAAATTAATTAATCTAATGAAATGAAATCA 9622
 QY 2521 TTTTCATATGAAATTAAGATATTAATTTAAAAAGTAATTCATGAGCAATTAATTAAG 2580
 Db 9623 TTTTCATATGAAATTAAGATATTAATTTAAAAAGTAATTCATGAGCAATTAATTAAG 9682
 QY 2581 TAGGATTTTCAATATGTTTAAAGATTAATTCAGGAAACAAAGTTTCTCAAAATTAAGCA 2640
 Db 9683 TAGGATTTTCAATATGTTTAAAGATTAATTCAGGAAACAAAGTTTCTCAAAATTAAGCA 9742
 QY 2641 GAAAAATTTTTCATGATATCAAGTCTTTTCATTTAAGTCTCTGAAATTAATCTGAT 2700
 Db 9743 GAAAAATTTTTCATGATATCAAGTCTTTTCATTTAAGTCTCTGAAATTAATCTGAT 9802
 QY 2701 TTCTAATTTATCAAGATCAAAAAATTAATTAATTAATTAATTAATTTTACTCA 2760
 Db 9803 TTCTAATTTATCAAGATCAAAAAATTAATTAATTAATTAATTAATTTTACTCA 9862
 QY 2761 ATGCTTAATTAATTAATTAATTAATTAATTTTCTGAGGAAACATATCAACTTT 2820
 Db 9863 ATGCTTAATTAATTAATTAATTAATTAATTTTCTGAGGAAACATATCAACTTT 9922
 QY 2821 TAAGTATGAAGTGAATTAATTAATTAATTTTCACTTAATTAATTTTAAAGTTTCAAT 2880
 Db 9923 TAAGTATGAAGTGAATTAATTAATTAATTTTCACTTAATTAATTTTAAAGTTTCAAT 9982
 QY 2881 ATAAAGATGAATTAATTAATTAATTTTCACTTAATTAATTTTAAAGTTTCAAT 2940
 Db 9983 ATAAAGATGAATTAATTAATTAATTTTCACTTAATTAATTTTAAAGTTTCAAT 10042
 QY 2941 ATGTCATATCAAGCTTTATTAATTCATGATTAATTCCTTTTAAAGATGCAATGAA 3000
 Db 10043 ATGTCATATCAAGCTTTATTAATTCATGATTAATTCCTTTTAAAGATGCAATGAA 10102
 QY 3001 TCAGCTCACTTCATCTGTAACAAATTAATCTGTTGTCATCTGTCAGACAGAGGTTT 3060
 Db 10103 TCAGCTCACTTCATCTGTAACAAATTAATCTGTTGTCATCTGTCAGACAGAGGTTT 10162
 QY 3061 ACCCTGACGACGAGATTCATCTGTCAGACAGAGCAATCAAGATCCGACGACACTTG 3120
 Db 10163 ACCCTGACGACGAGATTCATCTGTCAGACAGAGCAATCAAGATCCGACGACACTTG 10222
 QY 3121 TCTCATCAAGGTTGAATTAATTAAGCACTTGGAATTAAGTATTAAGATTAAGTCACTG 3180
 Db 10223 TCTCATCAAGGTTGAATTAATTAAGCACTTGGAATTAAGTATTAAGATTAAGTCACTG 10282
 QY 3181 TGGCAATGTTCTCTGACATGTCGAGAGCATTAACGTTTGCTTGCGTTTACTCAAAAGC 3240
 Db 10283 TGGCAATGTTCTCTGACATGTCGAGAGCATTAACGTTTGCTTGCGTTTACTCAAAAGC 10342
 QY 3241 AAAAGAAAGTAAAGAAAGAAAGTAAAGAAAGGAAAGAAAGTATTAAGATTTTAAAGC 3300
 Db 10343 AAAAGAAAGTAAAGAAAGAAAGTAAAGAAAGGAAAGAAAGTATTAAGATTTTAAAGC 10402
 QY 3301 ATGCAAAAAGTGAAGATCTCTGTTTAAATTAATTAATTAATTAATGCTGATGTTGCTGAGCA 3360

Db 10403 ATGCAAAAAGTGAAGATCTCTGTTTAAATTAATTAATTAATGCTGATGTTGCTGAGCA 10462
 QY 3361 GTGATCTGAATGAGAACAGCAGCAGCAAG 3390
 Db 10463 GTGATCTGAATGAGAACAGCAGCAGCAAG 10492

RESULT 2

BD233527

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 10492)
 Bass, J. J., Jeandlong, F., Kambadur, R. and Sharma, M.
 Novel promoter sequence of myostatin gene
 Patent: JP 2002519059-A 1 02-JUL-2002;
 AGRSEARCH LTD
 OS Bos sp. (bovine)
 PN JP 2002519059-A/1
 PD 02-JUL-2002 JP 2000558200
 PF 07-JUL-1999 JP 330902
 PR 07-JUL-1998 NZ
 PI JAMES J BASS, FERENC JEANDLONG, RAVI KAMBADUR, MRIDULA SHARMA PC
 C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/00, C12N5/ PC
 00

CC Novel promoter sequence of myostatin gene
 FH Key Location/Qualifiers
 FT promoter (1) . . (10402)
 FT enhancer (6728) . . (6734)
 FT enhancer (7874) . . (7880)
 FT micr signal (9819) . . (9832)
 FT TATA_ signal (9457) . . (9463)
 FT enhancer (9860) . . (9866)
 FT enhancer (10095) . . (10101)
 FT CAAT signal (10197) . . (10202)
 FT TATA_ signal (10240) . . (10247)
 FT TATA_ signal (10264) . . (10271)
 FT CDS Location/Qualifiers
 1. . 10492
 /organism="Bos sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:29061"

ORIGIN

Query Match

Best Local Similarity

Matches 3390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGCCTGCTCCGAGCAGATGAGAGGTTGCAATATCCCGCATGAGAGGAAAGAAACACTTC 60
 7103 TGCCTGCTCCGAGCAGATGAGAGGTTGCAATATCCCGCATGAGAGGAAAGAAACACTTC 7162
 QY 61 TTGAATTTTCTTGAGTACCTTTAATTCATTCATGATGTCGACCTCCGAGATTGATAGAG 120
 Db 7163 TTGAATTTTCTTGAGTACCTTTAATTCATTCATGATGTCGACCTCCGAGATTGATAGAG 7222
 QY 121 CTGACTCATTTATTCACAGTGTATGTTCTTACCAATCACTGCAACATGAATTAAGT 180
 Db 7223 CTGACTCATTTATTCACAGTGTATGTTCTTACCAATCACTGCAACATGAATTAAGT 7282
 QY 181 ATTCTGAAACACTGCCCCCTAGAGGAATCAAGGTTAAGTCCCTGACCTCTGATCA 240
 Db 7283 ATTCTGAAACACTGCCCCCTAGAGGAATCAAGGTTAAGTCCCTGACCTCTGATCA 7342

QY 241 CGTTTTGTTAACCATCAATTAATAACCTGTTGTTGTCGATTCGTTTTTAAATAT 300
| | | | |
Db 7343 CGTTTTGTTAACCATCAATTAATAACCTGTTGTTGTCGATTCGTTTTTAAATAT 7402
| | | | |
QY 301 CTTTATTTAATAGTACTGCTAATCTTCAACATTTGGTTCACACCAAAAGGCCCTAAT 360
| | | | |
Db 7403 CTTTATTTAATAGTACTGCTAATCTTCAACATTTGGTTCACACCAAAAGGCCCTAAT 7462
| | | | |
QY 361 AACTGGAAAGCCGTAATGAAGCTTACATTAACACATGTTTTTTCTATGAGAAATTT 420
| | | | |
Db 7463 AACTGGAAAGCCGTAATGAAGCTTACATTAACACATGTTTTTTCTATGAGAAATTT 7522
| | | | |
QY 421 TTCCTTCAGTCTGTCACAGCCCTTCTGTTAAATTTGTCGACAAATATATACATACATG 480
| | | | |
Db 7523 TTCCTTCAGTCTGTCACAGCCCTTCTGTTAAATTTGTCGACAAATATATACATACATG 7582
| | | | |
QY 481 AAATGACTATATTTAATCTGATTTTAACTGTACAGTTCAAGTTCAGTTCATTAAGTACCTTAC 540
| | | | |
Db 7583 AAATGACTATATTTAATCTGATTTTAACTGTACAGTTCAAGTTCAGTTCATTAAGTACCTTAC 7642
| | | | |
QY 541 CTGCTGTGGAACTATGATCAACATTCATCTAGAACTTTTGTATCTTCGAAATGAA 600
| | | | |
Db 7643 CTGCTGTGGAACTATGATCAACATTCATCTAGAACTTTTGTATCTTCGAAATGAA 7702
| | | | |
QY 601 AACTGTGACTATATTTGACAGTACGTTTCCCATCAACACCGACCCCTTCTCCAGCCCC 660
| | | | |
Db 7703 AACTGTGACTATATTTGACAGTACGTTTCCCATCAACACCGACCCCTTCTCCAGCCCC 7762
| | | | |
QY 661 CCGGACACACATCTCTTCTTCCATGACATGATCTCTGTGCTTGAGAAACACAGCCCTTCA 720
| | | | |
Db 7763 CCGGACACACATCTCTTCTTCCATGACATGATCTCTGTGCTTGAGAAACACAGCCCTTCA 7822
| | | | |
QY 721 ACACTAGCTTTGGGGGACACTGTAAGCAACAGATCACTCCCTACCGACCAAAATGAC 780
| | | | |
Db 7823 ACACTAGCTTTGGGGGACACTGTAAGCAACAGATCACTCCCTACCGACCAAAATGAC 7882
| | | | |
QY 781 ACAAAATATATAAAGCATGTGGCATATGCAATTTGCAAAAAGGCTGCTTGTGTAATGA 840
| | | | |
Db 7883 ACAAAATATATAAAGCATGTGGCATATGCAATTTGCAAAAAGGCTGCTTGTGTAATGA 7942
| | | | |
QY 841 GGGCTGAAAACAGGACAGAAATTTGATAGTTGACCTCAAGTGGGATCTGTGTGTTGA 900
| | | | |
Db 7943 GGGCTGAAAACAGGACAGAAATTTGATAGTTGACCTCAAGTGGGATCTGTGTGTTGA 8002
| | | | |
QY 901 AGCCTCAAAATTTTTCATGTTCTGTGACATACGCAAAATGCTTAAAGCACTGTAAAG 960
| | | | |
Db 8003 AGCCTCAAAATTTTTCATGTTCTGTGACATACGCAAAATGCTTAAAGCACTGTAAAG 8062
| | | | |
QY 961 ATTGATTATGAAAGTTAAGATTAATCTGACAAACATTAATGTGCAAGCAGGATCCAT 1020
| | | | |
Db 8063 ATTGATTATGAAAGTTAAGATTAATCTGACAAACATTAATGTGCAAGCAGGATCCAT 8122
| | | | |
QY 1021 GAATTAAGAGCACTGACATGTGGAATGATTAATCTTTGTTCTTTTATTCAGAGCGTA 1080
| | | | |
Db 8123 GAATTAAGAGCACTGACATGTGGAATGATTAATCTTTGTTCTTTTATTCAGAGCGTA 8182
| | | | |
QY 1081 AGAGAAAGGCTCACAAGGCTGCTTACACCATTTTACTAGAGAGCTAGGCTATGTCAG 1140
| | | | |
Db 8183 AGAGAAAGGCTCACAAGGCTGCTTACACCATTTTACTAGAGAGCTAGGCTATGTCAG 8242
| | | | |
QY 1141 TCGGTAGCTGGCAATTAACAACTGAGCAGTTCTAGTTCAATGAGAGATGAATTTAAC 1200
| | | | |
Db 8243 TCGGTAGCTGGCAATTAACAACTGAGCAGTTCTAGTTCAATGAGAGATGAATTTAAC 8302
| | | | |
QY 1201 ATATATCTCAACCCCTCTGCAATGAACAGAGACTTAAGTACAGTATACATGTC 1260
| | | | |
Db 8303 ATATATCTCAACCCCTCTGCAATGAACAGAGACTTAAGTATCAAGTATACATGTC 8362
| | | | |
QY 1261 ACTTACTATATGACAGGCACTGTAATCAACAAATTTACATGATTAATGAATTAAGTCC 1320
| | | | |
Db 8363 ACTTACTATATGACAGGCACTGTAATCAACAAATTTACATGATTAATGAATTAAGTCC 8422
| | | | |
QY 1321 CCAACTCTATGAGGAAGCTGAAGGTTAGAAAGTATCTCATTTATTACACAGTGG 1380
| | | | |

Db 8423 CCAACTCTATGAGGAAGCTGAAGGTTAGAAAGTATCTCATTTATTATACACAGTGG 8482
| | | | |
QY 1381 CAAACTGAGATCTGAATCTGAGTCTATCCAACTCCAGGACCTGAGATCCCAATTTGTACA 1440
| | | | |
Db 8483 CAAACTGAGATCTGAATCTGAGTCTATCCAACTCCAGGACCTGAGATCCCAATTTGTACA 8542
| | | | |
QY 1441 CAATTCATCAAGTTTAAAGGAAAGAAAGATTTGATTTGCTCAGAGTATAGGGGCA 1500
| | | | |
Db 8543 CAATTCATCAAGTTTAAAGGAAAGAAAGATTTGATTTGCTCAGAGTATAGGGGCA 8602
| | | | |
QY 1501 TATGTTCAATTTATACATTTACAAAGATTTATATGTTGAAAATTAATTTATCAACAAA 1560
| | | | |
Db 8603 TATGTTCAATTTATACATTTACAAAGATTTATATGTTGAAAATTAATTTATCAACAAA 8662
| | | | |
QY 1561 TAAACCTTAAAGCCGATCTAATACCTGCTCCGCAACAAAGACTATCTGAAATCCTTCA 1620
| | | | |
Db 8663 TAAACCTTAAAGCCGATCTAATACCTGCTCCGCAACAAAGACTATCTGAAATCCTTCA 8722
| | | | |
QY 1621 GGGCATCTGATTTGTGCTGCTGTTTCTTAACTCTTAATGATGAGCAAAATCTAATGCAAT 1680
| | | | |
Db 8723 GGGCATCTGATTTGTGCTGCTGTTTCTTAACTCTTAATGATGAGCAAAATCTAATGCAAT 8782
| | | | |
QY 1681 ATGTAAAGGCAATTTTCTCAAGAGATGATGATCTCTTAAGAAATTTGATGAATGCA 1740
| | | | |
Db 8783 ATGTAAAGGCAATTTTCTCAAGAGATGATGATCTCTTAAGAAATTTGATGAATGCA 8842
| | | | |
QY 1741 TTAACTTTGAGGCTACGATGATTTAGTCACTGAGGAGTAAATTAAGTGTACAA 1800
| | | | |
Db 8843 TTAACTTTGAGGCTACGATGATTTAGTCACTGAGGAGTAAATTAAGTGTACAA 8902
| | | | |
QY 1801 TGTGCGAAAGTATGACCTTAAATTAATTTGATATGAAACCACTGCACTCTTGTGG 1860
| | | | |
Db 8903 TGTGCGAAAGTATGACCTTAAATTAATTTGATATGAAACCACTGCACTCTTGTGG 8962
| | | | |
QY 1861 GAAAAAGTATGATTAATCTCTTTAGAGTCTTACCTTCCCAAAAGTATGAGGAAA 1920
| | | | |
Db 8963 GAAAAAGTATGATTAATCTCTTTAGAGTCTTACCTTCCCAAAAGTATGAGGAAA 9022
| | | | |
QY 1921 AATTAATCTGCTGAGGCTGAGAAACAGCTTCTGTTTCTGCTGCTAATTTGTTAGGT 1980
| | | | |
Db 9023 AATTAATCTGCTGAGGCTGAGAAACAGCTTCTGTTTCTGCTGCTAATTTGTTAGGT 9082
| | | | |
QY 1981 TTTTAATAGTTCATTTGATTAAGACCTTGTGCTCCCAAGCTAAGGTTGAGATTGATC 2040
| | | | |
Db 9083 TTTTAATAGTTCATTTGATTAAGACCTTGTGCTCCCAAGCTAAGGTTGAGATTGATC 9142
| | | | |
QY 2041 CTTACAGAGGCACTTCAATTTAGAGAAACAAAAAGCCCAATCTCTGCTCCAGACCTTA 2100
| | | | |
Db 9143 CTTACAGAGGCACTTCAATTTAGAGAAACAAAAAGCCCAATCTCTGCTCCAGACCTTA 9202
| | | | |
QY 2101 CCCCAAATCCCTGACAGGTGTGCTGCTCCGGTCAATAGAGAACTGGCAAGGAATAC 2160
| | | | |
Db 9203 CCCCAAATCCCTGACAGGTGTGCTGCTCCGGTCAATAGAGAACTGGCAAGGAATAC 9262
| | | | |
QY 2161 TAGAGGTGCAACGTACTAGAAAGTAAATAGACTAGACACTAAGTGAAGGAGA 2220
| | | | |
Db 9263 TAGAGGTGCAACGTACTAGAAAGTAAATAGACTAGACACTAAGTGAAGGAGA 9322
| | | | |
QY 2221 AAAATGGGCACTTCAATGATGTGTTCTTTCCCTTTCTGTGTTCACAAATGCTCCGATA 2280
| | | | |
Db 9323 AAAATGGGCACTTCAATGATGTGTTCTTTCCCTTTCTGTGTTCACAAATGCTCCGATA 9382
| | | | |
QY 2281 TAAATTAAGAGGTAATATACTAATTTTCTTTTCAACACTGGAAGGCTGAGAGAAA 2340
| | | | |
Db 9383 TAAATTAAGAGGTAATATACTAATTTTCTTTTCAACACTGGAAGGCTGAGAGAAA 9442
| | | | |
QY 2341 CTTTGTATACCATCATAAATTTCACTATCTTCTAAGTCAATCTATGTTATTTAAGTCA 2400
| | | | |
Db 9443 CTTTGTATACCATCATAAATTTCACTATCTTCTAAGTCAATCTATGTTATTTAAGTCA 9502
| | | | |
QY 2401 AATGCTGACATATCTCTTTTGTAAATTAACATGAAGAAACACATCTCTGACCAATAT 2460
| | | | |

* 1 24496: contig of 24496 bp in length
 * 24497 24596: gap of unknown length
 * 24597 47565: contig of 22970 bp in length
 * 47567 47665: gap of unknown length
 * 47667 192431: contig of 144765 bp in length.
 Location/Qualifiers

FEATURES

Source

1.192431
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-374H3"
 /clone_1db="RP42"
 1.24496
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"

misc_feature

24597..47566
 /note="assembly_fragment"
 47667..192431
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"

misc_feature

85439..192431
 /note="clone overlaps with GenBank Accession Number
 AC141852 clone RP42-269D20 (center project name djg)"

misc_feature

ORIGIN

Query Match 98.1%; Score 3327.2; DB 2; Length 192431;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 3371; Conservative 0; Mismatches 13; Indels 6; Gaps 3;

1 TGGCTGCTCCGAGCCATAGAGAGTTGCAATCCCGGACATGAGAGAGAAACACGCTTC 60
 105495 TGGCTGCTCCGAGCCATAGAGAGTTGCAATCCCGGACATGAGAGAGAAACACGCTTC 105436
 61 TTGAAATTTCTTGAGTACCTTCTTAATTCATTCATGCTGACCTCCGAGATTGATAGAG 120
 105435 TTGAAATTTCTTGAGTACCTTCTTAATTCATTCATGCTGACCTCCGAGATTGATAGAG 105376
 121 CTGACTCTCATTTATTCACAGAGGTTATGTTCTACCAATCACTGCAACATGAATTAAGTG 180
 105375 CTGACTCTCATTTATTCACAGAGGTTATGTTCTACCAATCACTGCAACATGAATTAAGTG 105316
 181 ATTCTGGAACCACTGCCCTAGAGGGAATCAAGAGTTAGATCCCGTCAAGCTCTGCTCA 240
 105315 ATTCTGGAACCACTGCCCTAGAGGGAATCAAGAGTTAGATCCCGTCAAGCTCTGCTCA 105256
 241 CGTTTTGTGAACCAATCAATTAATACCTTTGTTGTGTCATTTCTGTTTAAATAT 300
 105255 CGTTTTGTGAACCAATCAATTAATACCTTTGTTGTGTCATTTCTGTTTAAATAT 105196
 301 CTTTATTTAATAGATAGTGTGTAATTTCTCAACATTTGGTTCACAAACCAAGGCCCTATT 360
 105195 C-TTATTTAATAGATAGTGTGTAATTTCTCAACATTTGGTTCACAAACCAAGGCCCTATT 105142
 361 AACTGGAAGCCCTGAATGAAGCTTACATTAACAACATTTGTTTCTATGAAGAAAAATTT 420
 105141 AACTGGAAGCCCTGAATGAAGCTTACATTAACAACATTTGTTTCTATGAAGAAAAATTT 105082
 421 TTCTTCAGCTCCGCCACAGCCTTCTGCTTAAATTTGTGACAAATATATACATTAATG 480
 105081 TTCTTCAGCTCCGCCACAGCCTTCTGCTTAAATTTGTGACAAATATATACATTAATG 105022
 481 AAACGACTATTTAATAGATTTTAAGTATGATGACAGTTCAGTGCAATTAAGTACATTAC 540
 105021 AAACGACTATTTAATAGATTTTAAGTATGATGACAGTTCAGTGCAATTAAGTACATTAC 104962
 541 CTGCTGTGGAAGTATGATCAACATTCATCTATGAACCTTTTGTGATCTTCTCGAATTGAA 600
 104961 CTGCTGTGGAAGTATGATCAACATTCATCTATGAACCTTTTGTGATCTTCTCGAATTGAA 104902
 601 AACTGCACTATTTGACAGTAGCTTTCCCGCATCACCGCACCTTCTCTCAGGCCCC 660

104901 AACTGCACTATTTGACAGTAGCTTTCCCGCATCACCGCACCTTCTCTCAGGCCCC 104842
 661 CCGGACCAACATCTCTTCTTCCATGACATGATCTGTCGTAGAAACAGAGCCCTTCA 720
 104841 CCGGACCAACATCTCTTCTTCCATGACATGATCTGTCGTAGAAACAGAGCCCTTCA 104782
 721 ACACTAGCTTGGGGGCACTGTAGAGACAGATCACTCCCAACCAATATGAC 780
 104781 ACACTAGCTTGGGGGCACTGTAGAGACAGATCACTCCCAACCAATATGAC 104722
 781 ACAAAATATAAAAAGATGCTGCAATTTGATTTGCAAAAAGGCTGCTTCAATGTA 840
 104721 ACAAAATATAAAAAGATGCTGCAATTTGATTTGCAAAAAGGCTGCTTCAATGTA 104662
 841 GGGCTGAACAGAGCAGAGATTTGACTAGGTTGACCTGACCTGGATCTGTGTGGA 900
 104661 GGGCTGAACAGAGCAGAGATTTGACTAGGTTGACCTGACCTGGATCTGTGTGGA 104602
 901 AGCCTCAAAATTTTCCATTTGTTCTGTGATACGACAAATGCTTATAAAGCACTGTAAG 960
 104601 AGCCTCAAAATTTTCCATTTGTTCTGTGATACGACAAATGCTTATAAAGCACTGTAAG 104542
 961 ATTGATTTAGAGTAAATTAATCTCAGCAAGATTAATTTGCAAGCAGGATTCAT 1020
 104541 ATTGATTTAGAGTAAATTAATCTCAGCAAGATTAATTTGCAAGCAGGATTCAT 104482
 1021 GAATTAACGAGCACTGACATGTGGAATTAATCTTGTTCCTTATTCAGGCACTA 1080
 104481 GAATTAACGAGCACTGACATGTGGAATTAATCTTGTTCCTTATTCAGGCACTA 104422
 1081 AGAGAAAGGCTCAAGGCTGCTTACACCAATTTTACTAGAGAGCTAGCTATGTCAG 1140
 104421 AGAGAAAGGCTCAAGGCTGCTTACACCAATTTTACTAGAGAGCTAGCTATGTCAG 104362
 1141 TGGTAGCTGCAATTAACAATGAAGCACTTCTAGTCAATGAGAGATTTTAAAC 1200
 104361 TGGTAGCTGCAATTAACAATGAAGCACTTCTAGTCAATGAGAGATTTTAAAC 104302
 1201 ATATCTCAACCCCTCTGATGAAACAGAGCTAAGTACTCAAGTACAGTATCAGTC 1260
 104301 ATATCTCAACCCCTCTGATGAAACAGAGCTAAGTACTCAAGTACAGTATCAGTC 104242
 1261 ACTTACTATATGACAGGCACTGATCAACCAATTTACATGATTTATTAATGATCAGCC 1320
 104241 ACTTACTATATGACAGGCACTGATCAACCAATTTACATGATTTATTAATGATCAGCC 104182
 1321 CCAACACTCTATAGAGAGGTAAGGTTAGAGAAATCTCATTCATTAATTAACAGTGG 1380
 104181 CCAACACTCTATAGAGAGGTAAGGTTAGAGAAATCTCATTCATTAATTAACAGTGG 104122
 1381 CAAACTGAGATCTGAATCTAGTCTATCCAACTCCAGGACCTGAGATCCCAATTTGCTACA 1440
 104121 CAAACTGAGATCTGAATCTAGTCTATCCAACTCCAGGACCTGAGATCCCAATTTGCTACA 104062
 1441 CAATCTAATCAAGTTAAAGGAAAAAGATTTGATTTGCTCAGAGTATAGGGGCA 1500
 104061 CAATCTAATCAAGTTAAAGGAAAAAGATTTGATTTGCTCAGAGTATAGGGGCA 104002
 1501 TATGTTAACAATTAATCAATTAACAAGATTTATATGTTGAAAAATTAATTAACAACA 1560
 104001 TATGTTAACAATTAATCAATTAACAAGATTTATATGTTGAAAAATTAATTAACAACA 103942
 1561 TAAAACTTTAAGCCTGATCTATATCTGCTCCGCAACAAAGACTATCTGAATCTCTCA 1620
 103941 TAAAACTTTAAGCCTGATCTATATCTGCTCCGCAACAAAGACTATCTGAATCTCTCA 103882
 1621 GGGCATCTGGTTGTGTCTGGTTTCTTAAATCTTAAATGAGGCAATCTAATGCAAT 1680
 103881 GGGCATCTGGTTGTGTCTGGTTTCTTAAATCTTAAATGAGGCAATCTAATGCAAT 103822
 1681 ATGTAAGGCATTTTCTCAAGAGATGTAATCCTCTTAAATTTGATGTAAGAAATGCA 1740
 103821 ATGTAAGGCATTTTCTCAAGAGATGTAATCCTCTTAAATTTGATGTAAGAAATGCA 103762

QY	1741	TTAACTTTTCAGGCGTACAGTGTGACTTTTAGTGCCTGAGGCAAGTAAATTAAGTACAA	1800
Db	103761	TTAACTTTTCAGGCGTACAGTGTGACTTTTAGTGCCTGAGGCAAGTAAATTAAGTACAA	10370202
QY	1801	TGTGCGAAGTGTGACCTTAAAAATAATATTGTATGAACCACTGCACTCTCTTGG	1860
Db	103701	TGTGCGAAGTGTGACCTTAAAAATAATATTGTATGAACCACTGCACTCTCTTGG	10346420
QY	1861	GAATAAAGTAAAGTAACTCTCTTGAAGAGTCTTACGCTCCCAAAAGTAACTAGAAA	1920
Db	103641	GAATAAAGTAAAGTAACTCTCTTGAAGAGTCTTACGCTCCCAAAAGTAACTAGAAA	1035822
QY	1921	AATAAATCTCGTGGCCCTGAGAAACAGCTTCTGTCTTGCTGGCTATATTTGTTAAGT	1980
Db	103581	AATAAATCTCGTGGCCCTGAGAAACAGCTTCTGTCTTGCTGGCTATATTTGTTAAGT	10352222
QY	1961	TTTTATATAGTCAATTTGATTAAGCTTGTGGCTCCCAAAAGCTAAGTTGAGTTGATC	2040
Db	103521	TTTTATATAGTCAATTTGATTAAGCTTGTGGCTCCCAAAAGCTAAGTTGAGTTGATC	1034622
QY	2041	CCTACAGAGGCCACTTCAATTTAGAGAACAAAAGCCCAATCTGCTCCAGACCTTA	2100
Db	103461	CCTACAGAGGCCACTTCAATTTAGAGAACAAAAGCCCAATCTGCTCCAGACCTTA	10340202
QY	2101	CCCCAAATCCCTGCCAGGTGTCTGCCCTCCGCTCAATAGAGAACTGGCAAGAAAGTAC	2160
Db	103401	CCCCAAATCCCTGCCAGGTGTCTGCCCTCCGCTCAATAGAGAACTGGCAAGAAAGTAC	10334242
QY	2161	TAGAGAGTTCGCACATPACTAGAGAGTGAAGAAAATGGAATGACACATCTAGAGAAACAGA	2220
Db	103341	TAGAGAGTTCGCACATPACTAGAGAGTGAAGAAAATGGAATGACACATCTAGAGAAACAGA	10328222
QY	2221	AAAAATGGGCAACCTTCATGATGGTGTCCCTTCCCTTCTGTGTTCACATGCTCCGATA	2280
Db	103281	AAAAATGGGCAACCTTCATGATGGTGTCCCTTCCCTTCTGTGTTCACATGCTCCGATA	10322222
QY	2281	TAAATTTACAGAGGGTATGATTAACTAATTTTTCTTTACCACTGAGAGGCTGAGAAAA	2340
Db	103221	TAAATTTACAGAGGGTATGATTAACTAATTTTTCTTTACCACTGAGAGGCTGAGAAAA	10316222
QY	2341	CTTTCTTACCCCATCAATAAATTTACACATCTCTTCAAGTCAATCTATGTATTTCTAAGTCA	2400
Db	103161	CTTTCTTACCCCATCAATAAATTTACACATCTCTTCAAGTCAATCTATGTATTTCTAAGTCA	10310202
QY	2401	AATAGCTGACATATCTCTTGTATTAATAACAATGAAAAACAATCCTCTGACATAT	2460
Db	103101	AATAGCTGACATATCTCTTGTATTAATAACAATGAAAAACAATCCTCTGACATAT	10304242
QY	2461	TAAATCTGCACTTTAGAGATAGAGAAAGTAACTTAATCTAGTCAATTGAAACTGAAATACAA	2520
Db	103041	TAAATCTGCACTTTAGAGATAGAGAAAGTAACTTAATCTAGTCAATTGAAACTGAAATACAA	10298222
QY	2521	TTTTCATATGAAATAAAGATATTTAATTAAGTAATTCATGAGCAATTTAATTTAAG	2580
Db	102981	TTTTCATATGAAATAAAGATATTTAATTAAGTAATTCATGAGCAATTTAATTTAAG	10292222
QY	2581	TAGATATTTCAATATGTGTTAAGATTTATTCAGGAGAACAGTTCTCAATATTATAGCA	2640
Db	102921	TAGATATTTCAATATGTGTTAAGATTTATTCAGGAGAACAGTTCTCAATATTATAGCA	10286222
QY	2641	GAATAATCTTTACTAGTATCACAGCTTTTCAATTAAGCTTCTGTAATTAATCTGAT	2700
Db	102861	GAATAATCTTTACTAGTATCACAGCTTTTCAATTAAGCTTCTGTAATTAATCTGAT	10280222
QY	2701	TTCTAATTATACAGACTTAAAAATAATTTATATTAACAATTAATAATTTATTTACTTCAA	2760
Db	102801	TTCTAATTATACAGACTTAAAAATAATTTATATTAACAATTAATAATTTATTTACTTCAA	10274242
QY	2761	ATGCTTACTTAATATGTATTAATAATCAATTTATTTCTGAGGAGAAAGCATATCACTTT	2820
Db	102741	ATGCTTACTTAATATGTATTAATAATCAATTTATTTCTGAGGAGAAAGCATATCACTTT	10268222

QY	282	TAAGTATGAAGTGAATTAATTAAGATTTATTCACCTTAATTAATTTATTTAAAGTTTCACAT	2880
Db	102681	TAAGTATGAAGTGAATTAATTAAGATTTATTCACCTTAATTAATTTATTTAAAGTTTCACAT	1026
QY	2881	ATAAAGATGAATATAGATCTAAGTGTATATGTATATTTGTTAAATAAGATTTTAAATTTTTCG	2940
Db	102621	ATAAAGATGAATATAGATCTAAGTGTATATGTATATTTGTTAAATAAGATTTTAAATTTTTCGA	1025
QY	2941	ATGTCAATACAGCCTTTATTTATTTATTCATAGATTTATTCCTTTTAAAGAGTACTCAATGAA	3000
Db	102561	ATGTCAATACAGCCTTTATTTATTTATTCATAGATTTATTCCTTTTAAAGAGTACTCAATGAA	1025
QY	3001	TCAGCTCACCCCTTGACTGTATACAAATATACGTGTTGGTGACTTGGACAGACAGGGTTTAA	3060
Db	102501	TCAGCTCACCCCTTGACTGTATACAAATATACGTGTTGGTGACTTGGACAGACAGGGTTTAA	1024
QY	3061	ACCTCTGACAGCGAGATTCATTTGTGGAGCAAGAGCCATCACAGATCCCGACGACCTTG	3120
Db	102441	ACCTCTGACAGCGAGATTCATTTGTGGAGCAAGAGCCATCACAGATCCCGACGACCTTG	1023
QY	3121	TCTCATCAAAAGTTGGAATATATAAAAGCCACTTGGATATCAGTATATAAGATTCCTGTG	3180
Db	102381	TCTCATCAAAAGTTGGAATATATAAAAGCCACTTGGATATCAGTATATAAGATTCCTGTG	1023
QY	3181	TGGCAATGTGTCTCTCAGACTGGGACAGCAATTAACGTTTGGCTGGGCTTACTCAAAAGC	3240
Db	102321	TGGCAATGTGTCTCTCAGACTGGGACAGCAATTAACGTTTGGCTGGGCTTACTCAAAAGC	1022
QY	3241	AAAGAAAGATGAAGAGAGAGATGAAGACAAAGGGAAGAAAGTTGATTTGATTTTAAAGC	3300
Db	102261	AAAGAAAGATGAAGAGAGAGATGAAGACAAAGGGAAGAAAGTTGATTTGATTTTAAAGC	1022
QY	3301	ATGCAAAAACCTGCAAAATCTCTGTTTATATTTACCTATTTATGCTGATTTGTTGCTGCCCA	3360
Db	102201	ATGCAAAAACCTGCAAAATCTCTGTTTATATTTACCTATTTATGCTGATTTGTTGCTGCCCA	1021
QY	3361	GTGATCTGGAATGAAGAACAGGAGACAGAG 3330	
Db	102141	GTGATCTGGAATGAAGAACAGGAGACAGAG 102112	
RESULT 4			
AC141852/c			
LOCUS	233029 bp	DNA	linear
DEFINITION	Bos taurus clone RP42-269D20, WORKING DRAFT SEQUENCE, 4 ordered		
ACCESSION	AC141852		
VERSION	AC141852.2		
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.		
SOURCE	Bos taurus (cow)		
ORGANISM	Bos taurus		
REFERENCE			
AUTHORS	1 (bases 1 to 233029)		
	Antoniellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,		
	Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Carliaga,K.,		
	Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,		
	Gupta,J., Highghil,P., Han,J., Hansen,N., Ho,S.-I., Hu,P.,		
	Hutle,B., Idol,J.R., Karlins,E., Kwong,P., Latic,P., Lee-Lin,S.-Q.,		
	Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C.,		
	Markert,B., McDowell,J., Pagliarigan,C., Pearson,R., Portnov,M.B.,		
	Prasad,A., Reddi-Dugue,N., Schindler,K., Schueler,M.G., Shah,K.,		
	Sison,C., Stenitrop,S., Thomas,J.W., Thomas,P.J., Teipouri,V.,		
	Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.		
TITLE	NIH Comparative Sequencing Initiative		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 233029)		
AUTHORS	Green,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-MAR-2003) NIH Intramural Sequencing Center, 8717		
REFERENCE	Givemont Circle, Gaithersburg, MD 20877, USA		
	3 (bases 1 to 233029)		

AUTHORS
TITLE
JOURNAL

Green, E.D.
Direct Submission
Submitted (12-APR-2003) NIH Intramural Sequencing Center, 8717
Groveom Circle, Gaithersburg, MD 20877, USA
On Apr 12, 2003 this sequence version replaced gi:29124102.

COMMENT

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@hgri.nih.gov
----- Project Information
Center project name: djg
Center clone name: 269D20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies; the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 232253 bases at least Q40
Consensus quality: 232568 bases at least Q30
Consensus quality: 232684 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 232729; sum-of-contigs
Quality coverage: 11.69x in Q20 bases; agarose-fp
Quality coverage: 10.05x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
*
* 1 64168: contig of 64168 bp in length
* 64169 64268: gap of unknown length
* 64269 213649: contig of 149381 bp in length
* 213650 213749: gap of unknown length
* 213750 217486: contig of 3737 bp in length
* 217487 217586: gap of unknown length
* 217587 233029: contig of 15443 bp in length.

FEATURES

Source

1. .233029
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-269D20"
/clone_1db="RP42"
1. .107006
/note="clone overlaps with GenBank Accession Number AC141853 clone RP42-374H3 (center project name djp)"
1. .64168
/note="assembly_fragment
clone_end:Sp6
vector_side:left"
64269. .213649
/note="assembly_fragment"
213750. .217486
/note="assembly_fragment"
217587. .233029
misc_feature
misc_feature
misc_feature

ORIGIN

/note="assembly_fragment
clone_end:T7
vector_side:right"

Query Match 98.1%; Score 3327.2; DB 2; Length 233029;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3371; Conservative 0; Mismatches 13; Indels 6; Gaps 3;
ORIGIN
1 TGCTGCTCCGAGGCGATAGGAGGTGCAATCCCGGACACTGAGAGGAAACACGCTC 60
20057 TGCTGCTCCGAGGCGATAGGAGGTGCAATCCCGGACACTGAGAGGAAACACGCTC 19998
61 TTGAATTTCTTGAATCCTCTTAACTTCAATGCTGACCTCGAGATGATAGAG 120
19997 TTGAATTTCTTGAATCCTCTTAACTTCAATGCTGACCTCGAGATGATAGAG 19938
121 CTGACTCTCATTTATTCACAGGTGTTAGTTCTACCAATCATCTGCCAATGATTAAGTG 180
19937 CTGACTCTCATTTATTCACAGGTGTTAGTTCTACCAATCATCTGCCAATGATTAAGTG 19878
181 ATTCTGAAACACAGCCCTAGGAGGAACTCAAGGTTAGATTCCGTCAGCTGCTGCA 240
19877 ATTCTGAAACACAGCCCTAGGAGGAACTCAAGGTTAGATTCCGTCAGCTGCTGCA 19818
241 CGTTTGTGTTAACCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
19817 CGTTTGTGTTAACCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19758
301 CTTTATTTTAACTGACTGCTGCTAATTTCTTCAACATTTGCTTCAACACCAAGCCCTATT 360
19757 C-TTATTTTAACTGACTGCTGCTAATTTCTTCAACATTTGCTTCAACACCAAGCCCTATT 19704
361 AACTGAAGCCCTGATGAGAGCTTACATTAACACATGTTTCTTCAAGGAAATTAATTT 420
19703 AATACTGAAGCCCTGATGAGAGCTTACATTAACACATGTTTCTTCAAGGAAATTAATTT 19644
421 TTCTTCAAGCTCTGCAACAGCTTTCTTCAATTAATTAATTAATTAATTAATTAATTAAT 480
19643 TTCTTCAAGCTCTGCAACAGCTTTCTTCAATTAATTAATTAATTAATTAATTAATTAAT 19584
481 AACTGACTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
19583 AACTGACTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19524
541 CTGCTGAGAACTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
19523 CTGCTGAGAACTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19464
601 AACTGCTCATTTATTCAGAGTATTTCCCATTCACACCGCAACCTTCTCCAGGCC 660
19463 AACTGCTCATTTATTCAGAGTATTTCCCATTCACACCGCAACCTTCTCCAGGCC 19404
661 CCGGACACACATCTCTTCTTCCATGACAGTCACTCTGTGCTGAGAAACAGCCCTTCA 720
19403 CCGGACACACATCTCTTCTTCCATGACAGTCACTCTGTGCTGAGAAACAGCCCTTCA 19344
721 AACTAGAGCTTGGGGGAGCACTGTAAGCAAGATGCACTCCCTAGCGGCAACAAATGAC 780
19343 AACTAGAGCTTGGGGGAGCACTGTAAGCAAGATGCACTCCCTAGCGGCAACAAATGAC 19284
781 ACAAATAATTAATAAGCATGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 840
19283 ACAAATAATTAATAAGCATGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 19224
841 GGGCTGAACAAGGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
19223 GGGCTGAACAAGGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19164
901 AGCTCAAAATTTTCATGTTCTGTCATGCAAGCAAAATGCTTATTAAGCACTGTAGG 960
19163 AGCTCAAAATTTTCATGTTCTGTCATGCAAGCAAAATGCTTATTAAGCACTGTAGG 19104

Qy	961	ATTGATTTATGAGTTAAGTATTAATCTCAGCAAGCATPAATNGSCAGCAGGGATTCAT	1020
Dp	19103	ATTGATTTATGAGTTAAGTATTAATCTCAGCAAGCATPAATNGSCAGCAGGGATTCAT	19044
Qy	1021	GAATPAACGAGCACTGACCATGTGGAATAATCTTTGGTTTCTTTAATTCAGGCGAGTA	1080
Dp	19043	GAATPAACGAGCACTGACCATGTGGAATAATCTTTGGTTTCTTTAATTCAGGCGAGTA	18988
Qy	1081	AGGAAAGCGCTCAGAGGCTGCTTACACCAATTTTACTAGAGAGCTAGCTTAGTGCAG	1140
Dp	18983	AGGAAAGCGCTCAGAGGCTGCTTACACCAATTTTACTAGAGAGCTAGCTTAGTGCAG	18922
Qy	1141	TCCGTAGCTGGCAATTACAACTGACAGCTTCAATCTCAATGCGAGAGATGAATTTAAC	1200
Dp	18923	TCCGTAGCTGGCAATTACAACTGACAGCTTCAATCTCAATGCGAGAGATGAATTTAAC	18866
Qy	1201	ATAATCTCAACCCCTCTGCATGAAACAGAGACTPAAGTACTCAAGTACCAGTTATCAGTC	1260
Dp	18863	ATAATCTCAACCCCTCTGCATGAAACAGAGACTPAAGTACTCAAGTACCAGTTATCAGTC	18800
Qy	1261	ACTTACTATATGACAGGCACTGATCTCAACAAATTTACATGATTTATTTGAATTCATGCCC	1320
Dp	18803	ACTTACTATATGACAGGCACTGATCTCAACAAATTTACATGATTTATTTGAATTCATGCCC	18744
Qy	1321	CCAACACTCTATGAGGAAGCTGAAAGTTAGAGAATATCTCATTCATTATTTACACAGTGG	1380
Dp	18743	CCAACACTCTATGAGGAAGCTGAAAGTTAGAGAATATCTCATTCATTATTTACACAGTGG	18688
Qy	1381	CAAACTGAGATCTGAACTCAGGCTATCCAACTCCAGACCTGAGATCCCAATTTGCTACA	1440
Dp	18683	CAAACTGAGATCTGAACTCAGGCTATCCAACTCCAGACCTGAGATCCCAATTTGCTACA	18622
Qy	1441	CAATTCATATCAAGTTAAAGGGAAGAAAGATTTGATTTGCTCAGAAGTGATATAGGGCA	1500
Dp	18623	CAATTCATATCAAGTTAAAGGGAAGAAAGATTTGATTTGCTCAGAAGTGATATAGGGCA	18566
Qy	1501	TATGTGTACAAATTATPAACATTACAAAGATTTATATGTTGAAATAATTTATCAACAA	1560
Dp	18563	TATGTGTACAAATTATPAACATTACAAAGATTTATATGTTGAAATAATTTATCAACAA	18500
Qy	1561	TAAACCTTATPAACCTGATCTATATACGCTCGGCAACAAAGCTATCTGAAATCCCTCA	1620
Dp	18503	TAAACCTTATPAACCTGATCTATATACGCTCGGCAACAAAGCTATCTGAAATCCCTCA	18444
Qy	1621	GGGCACTCGTGTGTCGTGCTGTTTTCTCTTAATCTTTATATATGCGCAATCTATATGCA	1680
Dp	18443	GGGCACTCGTGTGTCGTGCTGTTTTCTCTTAATCTTTATATATGCGCAATCTATATGCA	18388
Qy	1681	ATGTPAAGGCCATTTTTTCTCAAGATGTGATACCTCTTATAGAAATTTGATGAATAATGCA	1740
Dp	18383	ATGTPAAGGCCATTTTTTCTCAAGATGTGATACCTCTTATAGAAATTTGATGAATAATGCA	18322
Qy	1741	TTAATCTTTACGGCTACGAGTGTGCAATTTTAGTGCATGAGGCGATAAATTAAGTATCA	1800
Dp	18323	TTAATCTTTACGGCTACGAGTGTGCAATTTTAGTGCATGAGGCGATAAATTAAGTATCA	18266
Qy	1801	TGTGCGAAAGTAGACCTTAATAAATAATTAATTTGATATGACACATGCACTCTCTTGGG	1860
Dp	18263	TGTGCGAAAGTAGACCTTAATAAATAATTAATTTGATATGACACATGCACTCTCTTGGG	18200
Qy	1861	GAATAAAGTATGAGATTACCTCTTAGAGATCCTTAGCTTCCCAAAAGTATGAGAAA	1920
Dp	18203	GAATAAAGTATGAGATTACCTCTTAGAGATCCTTAGCTTCCCAAAAGTATGAGAAA	18144
Qy	1921	AATTAATCTCCTGTGGCTGGAACACGCTTCTGTTTCTGTGCTGCTATATTTGTTTAGGT	1980
Dp	18143	AATTAATCTCCTGTGGCTGGAACACGCTTCTGTTTCTGTGCTGCTATATTTGTTTAGGT	18088
Qy	1981	TTTTTAATAGTTCATTTGATTTAGACCTTGTGGCTCCCAAAGCTATGAGAGTTTATC	2040
Dp	18083	TTTTTAATAGTTCATTTGATTTAGACCTTGTGGCTCCCAAAGCTATGAGAGTTTATC	18022
Qy	2041	CTTACAGAGGCCACTTCATTTAGAGAACAAAAGCCCATTTCTGTGCTCCACAGACTTA	2100

Db	18023	CCCTACAGAGGCCACTTCAATTTTAGAGAACAAAAGGCCCAATCTTCGCTCCAGACCTTA	17964
Qy	2101	CCCCAAATCCCTCCGACAGGTGTCTGCCCTCCGGTCAAAATGAGAACTGGCAAGGAATAC	2160
Db	17963	CCCCAAATCCCTGCCAGGTGTCTGCCCTCCGGTCAAAATGAGAACTGGCAAGGAATAC	17904
Qy	2161	TAGAGGTGCGACAGCTACTAGGAAGTGA AAAATGCACTGACACTACTGAGAGCAGA	2220
Db	17903	TAGAGGTGCGACAGCTACTAGGAAGTGA AAAATGCACTACTGAGAGAGAGA	17844
Qy	2221	AAAAAGGGGACCCCTTACATGATGCTGCTTCCCTTCGATGTTACAAATGCTCCGATA	2280
Db	17843	AAAAAGGGGACCCCTTACATGATGCTGCTTCCCTTCGATGTTACAAATGCTCCGATA	17784
Qy	2281	TAAATTACAGAGGGTAGATATACATATTTTTCTTTTACCACTGGAGGCTGAGAAAA	2340
Db	17783	TAAATTACAGAGGGTAGATATACATATTTTTCTTTTACCACTGGAGGCTGAGAAAA	17724
Qy	2341	CTTTGTTCACCATCAATAAATTCATATCTTCAAGTCATCTATGTTATTTCTAAGTCA	2400
Db	17723	CTTTGTTCACCATCAATAAATTCATATCTTCAAGTCATCTATGTTATTTCTAAGTCA	17664
Qy	2401	AATACCTGACATATTCCTCTTGTATA TAAACAATGAAAAACATCCCTGAGCAATAT	2460
Db	17663	AATACCTGACATATTCCTCTTGTATA TAAACAATGAAAAACATCCCTGAGCAATAT	17604
Qy	2461	TNATCTGCACTTTAGAGTAGAGAGATCTTAATACCTAGTCATTTGAACTGAAATCAA	2520
Db	17603	TNATCTGCACTTTAGAGTAGAGAGATCTTAATACCTAGTCATTTGAACTGAAATCAA	17544
Qy	2521	TTTTCATATGAATAAAGATATTTTAAAGTAATTCATGAGCAATTTAATTTAAG	2580
Db	17543	TTTTCATATGAATAAAGATATTTTAAAGTAATTCATGAGCAATTTAATTTAAG	17484
Qy	2581	TAGATTTTTCATTAATGTGTTAAGANTTTTACAGGAAACAAGTTTCTCAATTAATGCA	2640
Db	17483	TAGATTTTTCATTAATGTGTTAAGANTTTTACAGGAAACAAGTTTCTCAATTAATGCA	17424
Qy	2641	GAAATCTTTTCTAGATACAGATCTTTTCAATTTAAGCTTCTGAAATTAATCTGATTT	2700
Db	17423	GAAATCTTTTCTAGATACAGATCTTTTCAATTTAAGCTTCTCTGAAATTAATCTGATTT	17364
Qy	2701	TTCTAATTATACAGACTTAAATAATTAATTAATTAACAATAATAATTTATTTTACTCAA	2760
Db	17363	TTCTAATTATACAGACTTAAATAATTAATTAATTAACAATAATAATTTATTTTACTCAA	17304
Qy	2761	ATGCTTACTTAATATAGTATTAATAATCA TTTTATTTTCGAGGAAAGCATATCACTTTT	2820
Db	17303	ATGCTTACTTAATATAGTATTAATAATCA TTTTATTTTCGAGGAAAGCATATCACTTTT	17244
Qy	2821	TNAGTATGAAAGTGAATTAAGTTAATCACTTAATTAATTTTAAAGTTTCAAT	2880
Db	17243	TNAGTATGAAAGTGAATTAAGTTAATCACTTAATTAATTTTAAAGTTTCAAT	17184
Qy	2881	ATTAAGATGAATTAAGATCTAAGTGTATATGTTATTTTATAAAGTTTTTAATTTTGC	2940
Db	17183	ATTAAGATGAATTAAGATCTAAGTGTATATGTTATTTTATAAAGTTTTTAATTTTGC	17124
Qy	2941	ATGTCAATACAGCCCTTATTTATTCATAGATTTATTCCTTTTAAGAATGTCAATGAA	3000
Db	17123	ATGTCAATACAGCCCTTATTTATTCATAGATTTATTCCTTTTAAGAATGTCAATGAA	17064
Qy	3001	TCAGCTACCCCTTGACTGTAAACAAATACTGTTTGGTGA CTTGTCAGACAGGCTTTTA	3060
Db	17063	TCAGCTACCCCTTGACTGTAAACAAATACTGTTTGGTGA CTTGTCAGACAGGCTTTTA	17004
Qy	3061	ACCTCTGACAGCGAGATTCATTTGTGAGCAAGGCCAATCACAGATCCGACAGCACTTG	3120
Db	17003	ACCTCTGACAGCGAGATTCATTTGTGAGCAAGGCCAATCACAGATCCGACAGCACTTG	16944
Qy	3121	TCTCATCAAGTTGGATATTA AAAAGCCCTTGGATACAGTATTA AAGTTTCACTGCTG	3180

Db	16943	TCTCATCAAGGTGGAAATATPAAAAAGCCACTTGGAAATACAGTATPAAAAAGTTTACACTGCG	16883		
Qy	3181	TGGCAAGTTGTTCTCTCGACTGCGGACGCACTTAAAGCTTTGGCTGGCGTTACTCAAAAGC	3240		
Db	16883	TGGCAAGTTGTTCTCTCGACTGCGGACGCACTTAAAGCTTTGGCTGGCGTTACTCAAAAGC	16824		
Qy	3241	AAAAAGAAAGATPAAAAAGAAAGATPAAAAAGAAAGAAAGATPAAAAAGATPAAAAAGC	3300		
Db	16823	AAAAAGAAAGATPAAAAAGAAAGATPAAAAAGAAAGAAAGATPAAAAAGATPAAAAAGC	16764		
Qy	3301	ATGCAAAAGTCAAAATCTCTGTTATATTTACTTATTTATGCTGATTTGTTGCTGGCCCA	3360		
Db	16763	ATGCAAAAGTCAAAATCTCTGTTATATTTACTTATTTATGCTGATTTGTTGCTGGCCCA	16704		
Qy	3361	GTGATCTGAAATGAAAGACGACGACAGAG	3390		
Db	16703	GTGATCTGAAATGAAAGACGACGACAGAG	1674		
RESULT 5	AC140968/c	179158 bp	DNA	linear	HTG 03-APR-2003
LOCUS	AC140968/c	Sus scrofa clone RP44-467M16, WORKING DRAFT SEQUENCE, 3 ordered			
DEFINITION	piece.				
ACCESSION	AC140968				
VERSION	AC140968.2	GI:29501861			
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.				
SOURCE	Sus scrofa (p19)				
ORGANISM	Sus scrofa				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
AUTHORS	1 (bases 1 to 179158)				
	Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S. M., Benjamin, B., Blakeley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S., Carriaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Hashiguchi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E., Lario, P., Lee-Jin, S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B., Magulies, E. H., Masiello, C., Maekari, B., McDowell, J., Peguligan, C., Pearson, R., Portroy, M. E., Prasad, A., Reddik-Dugue, N., Schandler, K., Scheller, M. G., Sison, C., Stenitrop, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L., Wetherby, K. D., Wiggins, L., Young, A. and Green, E. D.				
TITLE	NISC Comparative Sequencing Initiative				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 179158)				
AUTHORS	Green, E. D.				
JOURNAL	Submitted (06-MAR-2003) NIH Intramural Sequencing Center, 8717				
	Grovenmont Circle, Gathersburg, MD 20877, USA				
REFERENCE	3 (bases 1 to 179158)				
AUTHORS	Green, E. D.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-APR-2003) NIH Intramural Sequencing Center, 8717				
	Grovenmont Circle, Gathersburg, MD 20877, USA				
COMMENT	On Apr 3, 2003 this sequence version replaced gi:28867000.				
	----- Genome Center				
	Center: NIH Intramural Sequencing Center				
	Center code: NISC				
	Web site: http://www.nisac.nih.gov				
	Contact: nisc_zoo@hgrl.nhl.gov				
	----- Project Information				
	Center project name: d1yq				
	Center clone name: 467M16				

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out

```

gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye 3, 100% of reads
Assembly program: Phrap; version 0.990139
Consensus quality: 178695 bases at least Q40
Consensus quality: 178865 bases at least Q30
Consensus quality: 178931 bases at least Q20
Insert size: 196000; agarose-fp
Insert size: 178950; sum-of-contigs
Quality coverage: 13.06x in Q20 bases; agarose-fp
Quality coverage: 14.30x in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
  1 93877: contig of 93877 bp in length
  * 93878 93977: gap of unknown length
  * 93978 105317: contig of 11340 bp in length
  * 105318 105417: gap of unknown length
  * 105418 179158: contig of 73741 bp in length.
    Location/Qualifiers
      1..179158
        /organism="Sus scrofa"
        /mol_type="genomic DNA"
        /db_xref="taxon:9823"
        /clone="RP4-467M16"
        /clone_1b="RP4"
      1..93877
        /note="assembly_fragment"
        clone_end:SP6
        vector_side:left"
      93978..105317
        /note="assembly_fragment"
      105418..179158
        /note="assembly_fragment"
        clone_end:T7
        vector_side:right"

ORIGIN
Query Match 55.8%; Score 1890.8; DB 2; Length 179158;
Best Local Similarity 79.1%; Prid. Match. 0;
Matches 2740; Conservative 0; GMMatches 532; Indels 190; Gaps 34;

  1 TGGTGCCTCCGAGCCATAGGAGAGTGGCAATCCGGGCACTGGAGAGGAAACACGTTCC 60
Db 65212 TCCTTCTCTGGGGCTTTGGAAATTTGCACTGACCTGAGAGAGGAAATTTAGGTTCC 651533
  61 TTGAAATTTCTTGAGTACCTCTTAATTCATTCATGCTGACCT-----CCGGAGATTGG 114
Db 65152 TCAAACTTTTAGTACCTATTAATCACTTCAATCTGACCTTGAGCACAAGTTGATTAG 650939
  115 ATAGAGCTGACTCATTTATTCACAGTGGTATGTTCTTACCAATCACTGCCAAC----- 169
Db 65092 ACTCAGTTGATCTCATTTATTCATTCACAGTGGTGTCTTATCAAAATCACTGCCAACACTGA 650333
  170 ATGAATTAAGGATTTCTGGAACCACTGCCCCCTAGGGGAACCTACAAAGTTAGATTCCCGTCA 229
Db 65032 ATGAATTAAGGATTCAGTCACTGAGCCATTCCTCTTGGGGGAAGATCAAAAGTTAGTTCCATATCA 649737
  230 GCCCTGGTACGTTTTTTGTTTACCAATCAATTAATTAACCTGTTTGTGTGCAATTTCTG 289
Db 64972 GCCTCTGGTACATTTTTCACCAACCAATCAATTAATTAACCTGTTTATGTTTCATTTCTG 649131
  290 TTTTAAATTAATCTTTAATTAATGATGCTAATTTCTTCAACATTTGGTTCAACACCA 349

```

Db 64912 TTTTAAATA--TTTATTTATATATATATTAATTAATTAACATTTGATTCATAGGCA 64855
Qy 350 AAGCCCTATTAATCGAAGCCCTGAATGAAGCTTACATACACACATTTGTTTCTATG 409
Db 64854 A-----CACACTATTAACCTGAGCTGAATTAATGCTTACATATATACATATTTTCCATTA 64799
Qy 410 AGAAAAATTTTCTTCACTGCTGCGCACAGCCTTCTGCTTAAATTTGGAACAAATA 469
Db 64798 AGAAATTAATTTCTCCATAGCTATAGCAACATCTTCTTTTAAATTTGTAATAATA 64739
Qy 470 TACATACATGAACCTGACTATTAATTAATGATTTTAACTGATCTGATGAGTTCAGTGCATTA 529
Db 64738 TACATAAGAC-----AGTTCAAGTGCAGTTCAAGTGCATTA 64702
Qy 530 AGTACATCTTACCTGCTGAGACATGATCAACATTCATCTAGAACTTTTGAATTT 589
Db 64701 AATACATTTATTTATTTGTTGCAATCATGATCAATTCATCTTCAAACTTTTGAACCTT 64642
Qy 590 CTCGAATTAATACTGCTGATCTATTTGACAGTAGAGCTTTCCCGATCACACCGACCTT 649
Db 64641 CCCAAATTTGAATAATTTGATCTATTTGAACAGTAGTGTTCCTCCATCCCAACCCACCTT 64582
Qy 650 CTTTCAGCCCCCGGACACCATCTTCTTTCCATGACATCTCTGCTGCTAGAAC 709
Db 64581 CTTCCACACCTGCTGAACACCATCTTCTTCCATGACACCTCTCTGATGTAAGAGC 64522
Qy 710 ACAGCCTTCAACCTAGCTTGGGGGGCACTGTAAGCAAGAGATCCTCCATACGCGC 769
Db 64521 ACAG-----CACTATGCTTTGGAGACATTTTAAACAGCAATCTCCCTA----- 64475
Qy 770 ACCAATGCAACAAAAATATATAAAGCATGTGCGATATGCAATTTGCAAAAGGCTGCT 829
Db 64474 --AAAAACAAGTAAATATGAAAAACATGTACTATACAGATTTGCAAAAGGGAACCTT 64417
Qy 830 GC-TAAGTATGAGGCTGAAAACAAGCAGAGAAATTGACTAGTTGACTCAGCTGGATC 888
Db 64416 GCTTAAATCTGAGAGCTGAAAACAAGCAGAGAAATTGCTTGTGACCTCACTGAGGACC 64357
Qy 889 CTGTGTGTGGAACCTCAAAATTTTCCATGTTGCTGTCATACCAACAAATGCTTATTA 948
Db 64356 CTGTGT-----CTTCAAAATTTTGTCTTTTGTATACAGCGCTTGAATGACTATCA 64305
Qy 949 AGCATGTAAAGATTTGATTAATGAATTAG--ATAATCTCAGCAAG--ACATAAATGTGA 1006
Db 64304 AGTCCATGATGATTTGATTTGAGATTAAATTAATTTTCAAGAAATGACAAATTTGCA 64245
Qy 1007 AGCAGGATTCATGAATTAACAGACACTGACCATGTGAAATGATTAATCTTTGTTCTT 1066
Db 64244 TACAGGAATTCATGAATTAATGACATTAAGCTGTACGGAATGA--CAGCTTTGTTCTT 64186
Qy 1067 TATTCAGGAGTAAAGGAAAGGCTCAGG----- 1039
Db 64185 TACTTCAGGATCTAAGAAATTAAGAGCTGAATATGCGCTTGTATAGGTTGTCTTGAAAT 64126
Qy 1100 -GCTGCTTACACCATTTTACTAGAGCTAGCTATGTCACTGCTGCTGCTGCAATTA 1158
Db 64125 ATTTCACTTGCTCAGAGGCTTAAAGAGCTAGCTATGTGAATGATTAATTTGAGAGTAC 64066
Qy 1159 AAATCTGAAGATTTCTAGTTCAATGAGAGATGAATTTAACCATTAATCTCAACCCCTCT 1218
Db 64065 AAATCTGAAGATTTCTAGTTCAATGAGAGATGAATTTAACCATTAATCTCAACCCCTCT 64006
Qy 1219 GCATGAAGAGAGCTAGTATCAAGTACAGATTAATCACTTAACTTAATATATAGAGC 1278
Db 64005 GCATGAAGAGAGAGCTAGTATCAAGTATCAAGTATTAATCACTTAACTTAATATAGAGC 63946
Qy 1279 ACTGTACTCAACAATTTATCATGTATTTATGAATTAATGATGAGGAA 1338
Db 63945 ACTATATCTCAGCATTTTATCATGATTAATTTGAATTAATGATGAGGAA 63887
Qy 1339 GCTGAAGTTTGAAGATATCTCATTCATTTATTAACAGTGGCAACTGAGATCTGAAT 1398
Db 63886 GAACTGAATATCTCATTCATTAATTAACATGTTAGAAAGTGGCAAGTTGATTTGAAT 63827
Qy 1399 CAGGTCTATCAACTCCAGAGACTGAGATCCCAATTTCTACACAATTTCAATCAATTA 1458
Db 63826 CAGGTCTATCTGACTCCAAAGCTT-ATTTCCCAAGACTTTGCAATTTCTAATCAAGTTTA 63768
Qy 1459 AAGGAAAAAGATTTGATTTGCTCAGAGTGTATGAGGCAATGTT--ACAATTA 1515
Db 63767 AA-----AAAAAAATCTGATTTACTCAGAGATGTATGAGAGATGTATGATTAATA 63710
Qy 1516 ACATTAAGAAATTTATATGTTGAATAATTAATTAATCAAAATTAATTAATTAATGAC 1575
Db 63709 ATATTTACAAAGATTTATATGTTGAATAATTAATTAATTAATTAATTAATTAATGAC 63653
Qy 1576 CTGATCTAATATCTGCTCCGCAACAAAGCTATCTGAATCTCTGAGGCACTGCTGTTGT 1635
Db 63552 CTGATCTAATATCTGCTCCGCAACAAAGATTTTGAATCTCTGAGGCACTGCTGTTGT 63593
Qy 1636 GTCTGTTTCTTAACTTTTAATGATGGCAATCTAATGATTAATGATTAAGCCATTTT 1695
Db 63592 GTCTGTTTCTTAACTTTTAATGATGAGCAATGATTAATGATTAATGATTAAGCCATTTT 63534
Qy 1696 TTCTCAAGATGATGATATCTCTTAAGAAATTTGATGAATAATGATTAATTTCAAGCT 1755
Db 63533 TTCTCAAGATGATGATATCTCTTAAGAAATTTGATGAATAATGATTAATTTCAAGCT 63474
Qy 1756 ACTGATTTGATTTTGAAGCACTGAGGAGTAAATTTGATGAATGATGAGCAAGTATG 1815
Db 63473 ACTGATTTGATTTTGAAGCACTGAGGAGTAAATTTGATGAATGATGAGCAAGTATG 63414
Qy 1816 ACCTAAAAATTAATTTGATTAATGAACCACTGACCTCTTTGG-----GAAAAAAG 1868
Db 63413 ACCTAAAAATTAATTTGATTAATGAACCACTGACCTCTTTGGAAAAAATTAATTAAT 63354
Qy 1869 TAAATGATTAATCTCTTAAGAGATCTTAAGTTTCCCAAAATGATGAGAAATTAATC 1928
Db 63353 TAAATGATTAATCTCTTAAGAGATCTTAAGTTTCCCAAAATGATGAGAAATTAATC 63294
Qy 1929 TCTGTGCTGGAAGACGCTTCTGTTCTGCTGCTGCTAATTTGTTAGGTTTAAATA 1988
Db 63293 TCTGTGCTGGAAGACGCTTCTGTTCTGCTGCTGCTAATTTGTTAGGTTTAAATA 63234
Qy 1989 GTTCATTTGATTAATGACCTTGTGCTCCCAAGCTAAGTTGAGATTTGATTCCTTACGA 2048
Db 63233 GTTCATTTGATTAATGATCTTGTGCTCCCAAGCTAAGTTGAGATTTGATTCCTTACGA 63174
Qy 2049 GGGCACTTCAATTTAGAGAAACAAAGCCCATCTCTGCTCCAGACCTTAACCCCAAT 2108
Db 63173 GGGCACTTCAATTTAGAGAAACAAAGCCCATCTCTGCTCCAGACCTTAACCCCAAT 63114
Qy 2109 CCGTGCAGAGTGTGCTGCTCCGCTCAATGAGAACTGGCAAGAGATCTAGAGGT 2168
Db 63113 CCGTGCAGAGTGTGCTGCTCCGCTCAATGAGAACTGGCAAGAGAGGTGCAAC----- 63059
Qy 2169 CGCAGATTAAGAGATGAGAAATGAGCTAGACATCTAGAAAGCAGAAATATGG 2228
Db 63058 -----CTAGACAGAAATTTGGAAACAGAAATATGG 63028
Qy 2229 CACCTTCATGATGTTGTTCTTCCCTTCTGCTGTTCACAATGCTCCGATTAATTTAC 2288
Db 63027 CACCTTCATGATGTTGTTCTTCCCTTCTGCTGTTCACAATGCTCCGATTAATTTAC 62968
Qy 2289 AGAGGTATGATTAATCAATTTTCTTTTACCTGGAAGGCTGAGAGAAATCTTGTTA 2348
Db 62967 AGAGGTATGATTAATCAATTTTCTTTTACCTGGAAGGCTGAGAGAAATCTTGTTA 62909
Qy 2349 CCAATCAATTAATTTCACTATC-----TTCTAATGCTATTTATG--TTATTTAAGATCAA 2401
Db 62908 TCAATCAATTAATTTCACTATC-----TTCTAATGCTATTTATG--TTATTTAAGATCAA 62849
Qy 2402 ATAGCTGACATATCTCTTTGATTAATAACAATGAAAAACATCTCTGAGCAATATT 2461
Db 62848 ATAGCTGACATATCTCTTTGATTAATAACAATGAAAAACATCTCTGAGCAATATT 62789

OY	2462	AATCGCAACTTTGGATAGGAAGTAACTTAATACAGCAATCTGAAATACAT	2521
Db	62788	AATTCGCACTTTAGGATAGGAAAAA-----TCAGTGGAAACCTGACAGAT	62740
OY	2522	TTTCATATGAATAAAGATATTTATTTAAAGTAATCCATGAGCAATTTATATTA-ANG	2580
Db	62739	TTTCACGGTAATAAAGATATTTATTTAAATATATTCATGTGTAATATACAGATTAAG	62680
OY	2581	TAGATTTTCTTATGTGTTAAGATTTATTCAGGGAAACAAGTTCTCAATATATAGA	2640
Db	62679	TATATATTTCTATATGTACTAGAAATTTAGTCAAGAAAACAGATTCTCAATATATAGT	62620
OY	2641	GAAATCTTTTACACAGATCACAGCTCTTTCAATTAAGT-----TTCCTGATPAATC	2694
Db	62619	GAAATATATTTTACTAGATCACATCTTAAATTTATTTACGCTCTCTTAATTTAATC	62560
OY	2695	TGTATTT-TCTAATATACAGACTAAAAATTAATTAATTAACAATAAAAATATTTT	2753
Db	62559	TGTATTTCTCGATTAACAGGACTAAAAATTTTAAACAGCAATBAATTTCTTTT	62500
OY	2754	ACTTCAATGCTTACTTAATATAGTATAAATCATTTTATTTCTGAGGAAAA--GCATA	2811
Db	62499	TCCCAATATGTTGTCTAAATATATGTAATTCATTTTATTTTGTAGGAAAAAGCAAT	62440
OY	2812	TCACTTTTATAGATAGAGTAAGTAATTAAGATTTATCTAATTAATTAATTTTAA	2871
Db	62439	TCACTTTTATAGATAGAGTA--AAGAAATCTATATTTAAATTAACA-TTTTAA	62384
OY	2872	GTTTCACATATAAAGATGAATAGATCAAGT--GTAATGTATTTGTTAAATTAAGTT	2928
Db	62383	GTTTCACATATAAAGATTAATTAATTTAAGTCAAGTTATATATTTGTTAACATAGAT	62324
OY	2929	TTAATTTTGGCATGTCAATACAGCCTTTATTTATCATAGATTTATTCCTTTAAGAG	2988
Db	62323	TTAATTTTCAAATGTCAATATCTTTATTTATTTGATATTTTCTTTTATGAAG	62264
OY	2989	TAGTCAATATGAATAGCGTCAACCCCTGTGCTGAACAAATACGTGTGGTGAAGCTGGACA	3048
Db	62283	TAGTCAATATGAATAGCTCAACCCCTGTGCTGAACAAATACGTGTGGTGAAGCTGTGACA	62204
OY	3049	GACAGGGTTTAACTCTGTACAGGAGATTCATTTGTGGACAAAGACCATCACAGATCC	3108
Db	62203	GACAGGGTTTAACTCTGTACAGGAGATTCATTTGTGGACAAAGACCATCACATGATCC	62144
OY	3109	CGACGACACTTGTCTCATCAAGTTGGAATATTAAGAACCACTTGGATATACGATTAATA	3168
Db	62143	TGACGACACTTGTCTCATCA--GTGGAATATTAAGAACCACTTGGATATACGATTAATA	62086
OY	3169	GATTCACGTGGTGGCAAGTTGTCTCACAATGGGACAGCAATTAAGTTGGCTTGGCG	3228
Db	62085	GATTCACGTGGTGGCAAGTTGTCTCACAATGGGACAGCAATTAAGTTGGCTTGGCG	62026
OY	3229	TTACTCAAAAGCAAAAGAAAGTAAAGGAAGATTAAGAACCAAGGAAAGATTTGAT	3288
Db	62025	TTACTCAAAAGC-----AAAAGTAAAGGAAGAAATTAAGAACCAAGGAAAGATTTGAT	61971
OY	3289	GATTTTAAACCATGCAAAAATCTGCAATCTCTGTTATATTTACTATATTAATGCTGAT	3348
Db	61970	GATTTTAAACCATGCAAAAATCTGCAATCTATGTTATATTTAACTGTTATGCTGAT	61911
OY	3349	GTTGCTGGCCAGATGATCTGAATGAAGAAACAGCAGACAGAG 3350	
Db	61910	GTTGCTGGTCCGATGATCTGAATGAAGAAACAGCAGACAGAG 61869	

RESULT 6				
AY208121	AY208121	167404 bp	DNA	linear
LOCUS				
DEFINITION	Sus scrofa myostatin gene, complete cds.			
ACCESSION	AY208121			
VERSION	AY208121.1	GI:34484364		
KEYWORDS	.			

ORIGIN	55.6%	Score 1883.2	DB 4	Length 167404
Query Match	55.6%	Score 1883.2	DB 4	Length 167404
Best Local Similarity	79.2%	Pred. No. 0		
Matches 2740	Conservative	0	Mismatches 533	Indels 187
			Gaps	35
QY	1	TGCGTCCGGAAGCATAGGAAGTGGCAATCCCGCGCTGGAGAAGGAACACGCTTC	60	
DB	78739	TCCTTCTCTGCGGCTTGGAAAGTTGACGTCACTGACGAGGAAGAAATTAAGTTTC	78798	
QY	61	TTGAAATTTCTTGAGTACCTCTTAATTCATTCATGCTGACCT-----CCGAGATTGG	114	
DB	78799	TCAAATCTTCTTGAGTACCTATTAATCACTCAATCACTGACCTTGGACAGATGATTAG	78856	
QY	115	ATGAGCTGACTCTCATTTATTCACAGTGGTATGTTTCAACCAATCACTGCCAAC----	169	
DB	78859	ACTCAGTTGATTCATATTATTCATGAGAGTGTGTTCTATCAATCACTGCCAACACTGA	78918	
QY	170	ATGAATAAGATATCCCGAACACCTGCCCTAGGGGACPAACAAGTTAGATTCCCGTCA	229	
DB	78919	ATGATATAGATGATCACTGACCACTTGTCTCTAAGGGAGTACAAAGTTAGTTCCATATG	78978	
QY	230	GCCCTGCTGACGTTTTTGTAAACAATCAATTAATTAACCTGTTTGTGTGTCATTTCTG	289	
DB	78979	GCCCTGCTGACATTTTCAACCAACAATCAATTAATTAACCTGTTTGTGTGTCATTTCTG	79038	
QY	290	TTTTTAATATCTTATTTAATAGTACGTAATCTTCAACATTTGGTTTGCACAACAA	349	
DB	79039	TTTTTAATAT- -TTATATTATATATTAATTAATGATTATTAACATTTGATTCATAGGCAA	79096	
QY	350	AAGGCGCTATTAATCTGGAAGCCCTGATGAGAGCTTACATPAACACATTTGTTTCTATG	409	
DB	79097	A-----CACACTATPAATCAAGCTGAAATTAATGCTTTCATATPAATCAATATTTTTCATTA	79155	
QY	410	AGGAAAAATTTTCTTCACTGCTGCAACAGCTTCTTGCTTAAATTTGTGACAAATA	469	

Db	79153	AGGAAATTATTTCCCTTAAGTACAGACAAACCTCTTTTAAAAATGTGGTAAATA	79212
Oy	470	TACATAACAGAAACGACTATTAATTAACGTATTTAACTGTACAGCTTCACTGCGCATTA	529
Db	79213	TACATAAGAC-----AGTACGTGTACAGTTCACTGCGCATTA	79249
Oy	530	AGTACACTTACCTTGGTGGAACTATGATCAACATTCATCTCTAAGCTTTTGTACTT	589
Db	79250	AATACATTATATGTGTGCAATCAGATCAATACATTCATCAACACTTTTGTACCTT	79309
Oy	590	CTCGAATTGAAAACCTGTGACTCTATTTGACAGTAGCTTTCCCCATCAACCGACCTT	649
Db	79310	CCGAAATTGAAAATTTGTATCTATTTGAACAGTAGGTTCCCCCATCCACCCACCTCT	79369
Oy	650	CTTCACGCCCCCGGACCAACCATCTTCTTTTCATGACAGTATCTGTGCTAGAAAC	709
Db	79370	CCTCCACACACTGGTATCAACACATTCCTCTTCCATCAACAGCTTCTGTATGTAGAAC	79429
Oy	710	ACAGCCCTTCAACATACGTTTGGGGGCACTGTAAACACAGATCACTCCCTACCGCC	769
Db	79430	ACAG-----CACTATGCTTTGGGACATTTTAAACAGCAAAATCATCCCTTA-----	79476
Oy	770	ACCAATGTGACACAAAATATTAATAAGACTGTGGCATATCATGTCAAAAAGGGTGCTT	829
Db	79477	--AAAAACAAGTAAAAATGTGAAAAACATGTGATCTATACAGTTGCCAAAGAGGACCTT	79534
Oy	830	GC-TAAGTATAGGGCTGAAACAAAGCAGAAATGACATAAGTTGACCTCAGCTGGATC	888
Db	79535	GCTTAATCTGAGAGCTGAAACAAAGCAGAAATGCTGTGTAACCTCACTCACTGGGACC	79594
Oy	889	CTGTGTGTGGAAAGCTCAAAATTTTCCATGTTCTGTGCAATGCACAAAATGTTTAA	948
Db	79595	CTGTGT-----CCTCAAGATTTTGTGCTTTTGTACAGGACTTTGAAATGACATATCA	79646
Oy	949	AGCACTGAAAGATTTGATTAATGAAGTTAAG-ATTAATCTCAGCAG-ACATAAATGTGA	1006
Db	79647	AGTCCCATGAGTATTTGATTTCTGAGATTAATAATTTACAGAAAGTACGAAATTTGCA	79706
Oy	1007	AGCAGGGATTCATGAATTAACGAGCAGTGAACATGTGGAATGATATCTTTGTTCTT	1066
Db	79707	TACAGGGAATCCATGAATTAACGAGCATTAAGCTGTACGAAATGA-CAGCTTTGTTCTT	79765
Oy	1067	TATTCGAGGACGTAGAGAAAGCGCTCACAG-----	1099
Db	79766	TACTCCAGGTACTAAGATTAAGAGCTGAATGTGGCTGTGTGATAGGGTGTCTTGGAA	79825
Oy	1100	-GCGGCTTACACCATTTTACTAGAGGCTAGGCTATGTCAGTCGGTACGTGGCAATTAC	1158
Db	79826	ATTTCACCTTGCTTAGAGGCGCTAGAGGCTAGCGCTATGTGAACTGAATTTGGCAGCTAC	79885
Oy	1159	AAACTGAAGAGTTCTAGTTCATGTGAGAGTAAATTTAACATAATCTCAACCCCTCT	1218
Db	79886	AACTGAAAGAGTTCTAGTTCAATGTGAGAGTAAACTTAAGCATTTATCTCAAAACCCCTCT	79945
Oy	1219	GCATGAACAGAGCTAAGTACTCAAGTACAGTTTCACTCACTTAATATGACAGCC	1278
Db	79946	GCATGAACAAAGACCAAACTCAAGTACTAAGTATCAATCACTTAATATGACAGCC	80005
Oy	1279	ACTGTACTCAACATTTTCACTATTTATTTGAATTAACATGCCCCCAACCTATATGAGAA	1338
Db	80006	ACTATACTCAGCAATTTTACATGCAATTTTAAATTTACAT-CCCCCAAACTGAAGGTTTGA	80064
Oy	1339	GCTAAGAGTTTGAAGATGATCTCAATTCATTAATTAACAGTGGCAAACTGAGTCTGAAT	1398
Db	80065	GAACTAAGTATCTCATCCATTAATTAACATGTTTGAAGTGGCAAACTGAGTCTGAAT	80124
Oy	1399	CAGGTCTATCTGACTCCAAAGCCT-ATTCCCAACAGCTTTGCAATTTCTATTAAGTTTA	1458
Db	80125	CAGGTCTATCTGACTCCAAAGCCT-ATTCCCAACAGCTTTGCAATTTCTATTAAGTTTA	80183
Oy	1459	AAAGGAAAAAGATTTTGAATTTGCTCAGAAAGTGTATAGGGGCAATGTGTT--ACAATTA	1515
Db	80184	AA-AAAAAAATATCTGATTTTATCTCAGAAAGTGTATAGGCAATGTGTATGATTAATTA	80242

OY	1516	ACATTAAGAAGATTATATATGTAAATAAATTTATCAACAACAAATPAATACTTTATAC	1575
Db	80243	ATATTACAAAGATTATATGTAAATAAATTTATCA---AAAAACCCTTATPAC	802398
OY	1576	CTGATCTAAATACGCTCCGCAACAAAGCTATCTGAATCTTTCAGGGCATCTGTTGT	1635
Db	80300	CTGATCTAAATATGCTCCACAAACAAAGAAATTTCTGAATCTTTCAGGGCATCTGTTGT	80358
OY	1636	GTCTGTTTTCCCTTAATCTTTAATGATGGGCAAACTTAATGCAATTAATGAAGCCATTTT	1695
Db	80360	GTCTGTTTTCCCTTAATCTTTAATGATGGGCAAGCTTAATGCAATTAATGAAGCCCA-TTT	80418
OY	1696	TTCTCAAGAGATGTAGATACCTCTTAAGAATTTGATGAATAATGCAATTAATCTTTCAGGCT	1755
Db	80419	TTCTCAAGAGATGTAGATACCTCTTAAGAATTTGATGAATAATGCAATTAATCTTTCAGGCT	80478
OY	1756	ACTGAGTTGCATTTTAGTGCACTGAGGCGAGTAATTAATGTATCATGTGCCAAATGATGTC	1815
Db	80479	ACTGAGTTGCATTTTAGTGCACTGAGGCGAGTAATTAATGTATCAATGTGCCAAAGATGTC	80538
OY	1816	ACCTAAATAATTAATATTTGATATGAACCATGCACTCTCTTGG----GGAATAAGTAA	1871
Db	80539	ACCTAAATAATTAATATTTGATATGAACCATGCACTCTCTTGGATAATTAATTAATGTTA	80598
OY	1872	TGATTAACCTCTCTTGGAGTCCCTTAGCTTCCCAAAAGTAGTAGGAAATAATATCTCC	1931
Db	80599	TGATTAACCTCTCTTGGAGTCCCTTAGCTTCCCAAAAGTAGTAGGAAATAATATCTCC	80658
OY	1932	TGTGGCCCTGGAAACAGCTTCTGTTTCTGTGTGGCTATTTGTTTAGTTTAAATAGTT	1991
Db	80659	TGTGGCCCTGGAAACAGCTTCTGTTTCTCATGTGCATATGTTTATGTTAGCTCTTAAATAGTT	80718
OY	1992	CATTGATTAAGACCTTGTTGGCTCCCAAGCTAAGGTTGAGTTGATCCCTACAGAGCC	2051
Db	80719	CATTGATTAAGATCTTGTTGGCTCTTAAGGTTAAGTTGAGATTTGAGCTCTACAGAGCC	80778
OY	2052	CACTTCAATTTAGAAACAATAAGCCCATTTCTGTCTCCAGACCTTACCCCAAT-CC	2110
Db	80779	CACTTAAATTTAGAAACAATAAGCTCTATTTCTGTCTCCAGACCTTACCCCAATGTCC	80838
OY	2111	CTGCGAGGTCTGCCCCCGGCTCAATGTAGAACTGGCAAAAGAAAGTACTAGAGGTGCTG	2170
Db	80839	CTGCGAGGTCTGCCCCCTGTGTCAATGTAGAACTGGCAAAAGGGTGTGAAC-----	80899
OY	2171	CACAGTACTAGGAAGTAGAAAAATGACCTAGCACTACTGAGAGCAGAAAAATGCGCA	2230
Db	80892	-----CTAGCAACAATTTGGGAAACAAAAAATTTGGCA	80924
OY	2231	CCCTTCATGATGTTGCCCTTCCCTTCTGTGTTCAAAATGTCGGATATATTAACG	2290
Db	80925	CCCTTCATGATGTTGCCCTTCTCTTTATGTTTCAAAATTTGGCATATATTAACG	80988
OY	2291	AGGATAGATACATATTTTTTCTTTTACCACTGGAGGCTGAGGAAAACTTGTATACC	2350
Db	80985	AGATATAT-ACATACATTTTTTACTTTTCAACACTGGAAATCTGAGGCAAACTGCATATAC	81043
OY	2351	CATCATTAATTTCACTATC-----TTCTAAGTCATTTATG-TTATTTCTAAGTCAAT	2403
Db	81044	AGCATATTAATTTCACTATCTTTCTATATCTAAGTTATCTAAGTTATTTCTAAGTCAAGT	81103
OY	2404	AGCTGACAAATATCTCTTTGTATTAACAATGAATAACAAATCCTCTAGCAATTTTAA	2463
Db	81104	AGCTGACATATATCTCTTGTATTAACAATGAATAACAAATCTTCTAGCAATTTTAA	81166
OY	2464	TCTGCAACTTTAGATAGGAAGTAACTAATCTAGTCAATTGAACTGAATATCAATTT	2523
Db	81164	TCTGCAACTTTAGATAGGAATA-----TCAGTTGGAATCTAGACCAATTT	81212
OY	2524	TCATATGAATTAAGAATATTAATTAAGAATTCATGATAGCAATTAATATATTA-AGTA	2582
Db	81213	TCATGTAAATTAAGAATATTAATTAAGAATATTTCAATGTATATTAACAGATTAAGTA	81277

OY	17933	GTGTCAATGATGCGAAA---GTAGTACCTAAATAATTAATTTAGATGAACACCTGG	1849
Db	112274	GTGTCAATGTGCAAAAGTGTGATGACCTTAATAATAATTTATATAGCACCTGC	1122155
OY	1850	ACTCTCTTGG---GGAAAAAAAGTAATGATTAAGTCTCTTAGAGTCTCTAGCTTCCCC	1905
Db	112214	ATTCTCTGGGGGTGGGGGAAAGTAATGATTAATCTCTCTAGAGATCTCTTAGCTTCCCA	1121555
OY	1906	AAAATAGTAGGAAAAAATAATCTCTGTGGGCTGGAAAAGGTTCTGTTTCTGTGGTGG	1965
Db	112154	AAAACAGTAGGAAAGAAATCTCTGTGGCTGGAAAACAGT--TGTTTCACTGGC	1120997
OY	1966	TATATTTGTTTGGTTTTTAAATAGTTCATTTTGTATGATAGCTTGTGGCTCCCAAGCTAAG	2025
Db	112096	TATGTTTGCTTAGCTCTTTAAACGTTCAITTTATGATTTAGTCTGTGGCTCCCAAGCTAAG	112037
OY	2026	GTTTGAGATTTGATCCCTTACAGAGGCACTTCAITTTAGAGAAACAAAAGCCCATTTCTC	2085
Db	112036	GTTTGAGATTTGATCCCTTACAGAGGCACTTCAITTTAGAGAAACAAAAGCTCCATTTCTC	1119777
OY	2086	TGCTCCCAAGACCTTACCCCAAAATCCCTGCGAGGTGTGTCGCCCTCGGATC-AAATAGAAA	2144
Db	111976	TGCTCCCAAGACCTTACCCCAAAATCCCTGCGAGGTGTGTCGCCCTCGATCAAAATAGAAA	1119177
OY	2145	CTGGCAAAAGAAAGTACTAGAGAGTGCACAGTACTAGAAATGAAAAATGACTAGCAC	2204
Db	111916	CTGGCAAAAGG-----GTGCAAAACCAATGCTAG	1118889
OY	2205	ACTACTGAGAAACAGAAAAATGGGCAACCTTCAAGATGATGTTCCCTTCTGTGTG	2264
Db	111888	TATTTGGGCAATGTTAAAAAGCAACCCCTTTATATGATGCTTTTCACTTTTATGTGC	1118299
OY	2265	TCACAATGCTCCGATTAATTAATTACAGAGGATGATTAATCAATTTTCTTTTACACCT	2324
Db	111828	TGACAGTATTTGATTAATTAATTACATGATGAT--ACTAGCTTTTACTTTTGGCAAT	1117700
OY	2325	GGAAGGCTGAGAAAAACTTG-----TTACCATATATAAATTCATCTATCTT	2371
Db	111769	GGAATATCTGAGGCAAACTGTGCTTGGTGCATGCTGCAATGCAAGTGAACCAATGCTCTT	1117100
OY	2372	CTAAGTCATTCATGTATTTCTTAAGATCAAAATGCTGACAAATTCCTCTTTGTAAATAC	2431
Db	111709	CTTATTC--TGAGTTCAGAGATTTTTCATACGTGACATTAATCTCTTGGTAATAC	1116522
OY	2432	AATGAAAAAACATCTCTGAGCAATTAATCTGCAACTTTAGATAGAAAGTAACCTT	2491
Db	111651	AATGAAAGAAACATCTCTGATTAATTAATCTGCACTTTACAAGAGAAAGAACTG	1115922
OY	2492	AATCTAGTCATTTGAACTGAATATCAATTTTCAATATAATTA---GATATATTT	2546
Db	111551	AATAGTATTCACCCAAAAATGAGCAACATTTTCATAGTAAGAAAAACATTAATTAAT	1115522
OY	2547	TAAAAATTAATTCATGAGCAATTTAATTTAAAGTA---GATTTTCAATATGTGTTAAG	2603
Db	111531	TAAAAATTAATTCATGTGTAATTAATCAATTAATTAATGATTTTAAAGTGTCTAAG	1114722
OY	2604	AATTTATTCAGGAAACAAAGTTTCTTAATTAATAGCAAAAAATCTTTACTAGTATAC-	2662
Db	111471	AATTTAAACAGGAAAAACAAGCTCTCAAAATTTTGTGCAAACTTTTACTGTATACACT	1114122
OY	2663	-----AGCTTTTCAATTTAAAGTCTCTGAAATTAATCTGTATTTTCTAATTTAC	2713
Db	111411	ATCTTAGCTTTCTTTAATTTCAAGGCTTCTCTAATTAATTCATATTTCCCTAATTAAC	1113532
OY	2714	AGAC--TAAAAATAATTTAATTAACAATAAATAATTAATTTTACTTCAATGCTTACTT	2770
Db	111351	CGGCACTTAAACATTAATTTAAATACAGTAATATAAATAATTAATTTTCAATGCTGCT	1112922
OY	2771	AAATAGTAAATAATCATTTTATTTTGTGAGGAAAAAGCATCACTTTTAAAGTATGA	2830
Db	111231	AAATTAACAGATCTCATTTTAAATTTTAAAGAGATACATTTTCACTTTTAAAGCATGA	1112322

QY	2831	GGGT-AAATTAGATTTCATCTTAAATTAATTTTAAAGTTTCACATTAAGAAG	2889
Db	111231	GGTAAACTAGATTCTACTTATTTAAATTAACCT--GAGTTTCACATGAAGATT	111175
QY	2890	AATAAGATCTAAGTGTATATGTTATGTTAAATA-----AGCTTTTAAATTTTCGATG	2943
Db	111174	AAGAAGATTAAAGTAGTTCTATATTAATGTTAACATAGACTTTATTTTTTTCAAATG	111155
QY	2944	TCACTACAGCCTTATATTCATAGATTATTCCTTTTAAAGATAGTCAATGAATCA	3003
Db	111114	TCACATATATCTTTCATATATTTGTAAGATTATTTCTTTATAGATGCGAATGATCA	111055
QY	3004	GCTACCCCTGACCTGTAACAAAATACTGTTTGTGACTGTGACAGACAGGGTTTAAAC	3063
Db	111054	GCTACCCCTTGACTGTAACAAAATACTGTTTGTGACTGTGACAGACAGGGTTTAAAC	110995
QY	3064	TCTGACAGCGAGATTCATTTGTGAGCAAGAGCCAAATCAAGATCCGACGACACTGTCT	3123
Db	110994	TCTGACAGCGAGATTCATTTGTGAGCAAGAGCCAAATCAAGATCCGACGACACTGTCT	110935
QY	3124	CATCAAGTTGGAATATATATAAAAGCCACTTGGAAATCAGTATATAAAGATTCACCTGGTGG	3183
Db	110934	CATCAAGTTGGAATATATATAAAAGCCACTTGGAAATCAGTATATAAAGATTCACCTGGTGG	110875
QY	3184	CAAGTGTCTCTCAACACTGGGCGAGCATTAAGTTGGCTTGGCGTTACTCAAAAGCAAA	3243
Db	110874	CAAGTGTCTCTCAACACTGGGCGAGCATTAAGTTGGCTTGGCGTTACTCAAAAGCAAA	110815
QY	3244	AGAAAGATTAAGGAAGAGTAAAGAACAAAGGAAAAAGATTGATTGATTTTAAAAACATG	3303
Db	110814	AGAAAGATTAAGGAAGAGTAAAGAACAAAGGAAAAAGATTGATTGATTTTAAAAACATG	110755
QY	3304	CAAAAACCTGCAAAATCTCTGTTTAAATTTAACTATTTAAGTGAATGTTGCTGGCCAGTG	3363
Db	110754	CAGAAACCTGCAAAATCTATGTTTAAATTTTAACTGTTATGCTGATTTGCTGCAGTG	110695
QY	3364	GATCTGAATGAGAACAGCGACGAGAA 3389	
Db	110694	GATCTGAATGAGAACGTCGACGAGAA 110669	

RESULT 8				
AC073120/c				
LOCUS	AC073120	43877 bp	DNA	linear PRI 07-NOV-2001
DEFINITION	Homo sapiens BAC clone RP11-612D17 from 2,			complete sequence.
ACCESSION	AC073120			
VERSION	AC073120.5			
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 43877) Sulston,J.E. and Waterston,R. Toward a complete human genome sequence Genome Res. 8 (11), 1097-1108 (1998) 99063792 MEDLINE PUBMED 9847074 2 (bases 1 to 43877) Nguyen,C., Maupin,R., Dignan,G. and Wedgeworth,P. The sequence of Homo sapiens BAC clone RP11-612D17 unpublished 3 (bases 1 to 43877) Waterston,R.H. Direct Submission Submitted (08-JUN-2000) Genome Sequencing Center, Washington MO 63108, USA 4 (bases 1 to 43877) Waterston,R.H. Direct Submission Submitted (18-Sep-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis,			
TITLE				
AUTHORS				
JOURNAL				

REFERENCE MO 63108, USA
 5 (bases 1 to 43877)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Sep 18, 2001 this sequence version replaced gi:114043001.

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: MUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0612D17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPI1-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanesse, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6
 NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RPI1-16H9; the clone sequenced to the right is RPI1-128011, 200 bp overlap. Actual start of this clone is at base position 31717 of RPI1-16H9; actual end is at base position 114077 of RPI1-128011.

Data from AC011941 was used to finish this clone.

FEATURES

Source

1. 43877
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-612D17"
 /clone_11b="RP11-11"
 repeat_region
 327..350
 /rpt_family="AT_rich"
 repeat_region
 1135..1211
 /rpt_family="L1"
 repeat_region
 1198..1234
 /rpt_family="L1"
 repeat_region
 1654..1679
 /rpt_family="AT_rich"
 repeat_region
 1682..1999
 /rpt_family="(A)n"
 repeat_region
 1976..2008
 /rpt_family="Alu"
 repeat_region
 1976..2008
 /rpt_family="(A)n"

repeat_region 2020..2163
 /rpt_family="L1"
 repeat_region 2175..2277
 /rpt_family="L1"
 repeat_region 2278..2562
 /rpt_family="Alu"
 repeat_region 2563..2623
 /rpt_family="L1"
 repeat_region 2633..2720
 /rpt_family="ERV1"
 repeat_region 2721..3136
 /rpt_family="L1"
 repeat_region 5998..6019
 /rpt_family="(TTTC)n"
 repeat_region 6021..6196
 /rpt_family="(TTC)n"
 repeat_region 6189..6496
 /rpt_family="Alu"
 repeat_region 6576..6626
 /rpt_family="MIR"
 repeat_region 9551..9579
 /rpt_family="AT_rich"
 repeat_region 9783..9829
 /rpt_family="MIR"
 repeat_region 13375..13519
 /rpt_family="MIR"
 repeat_region 13619..14081
 /rpt_family="MER1_type"
 repeat_region 14106..14413
 /rpt_family="L2"
 repeat_region 15315..15662
 /rpt_family="ERV1"
 repeat_region 15723..15868
 /rpt_family="ERV1"
 repeat_region 15922..16087
 /rpt_family="L1"
 repeat_region 16029..16085
 /rpt_family="AT_rich"
 repeat_region 16325..16624
 /rpt_family="L1"
 repeat_region 16894..16925
 /rpt_family="MALR"
 repeat_region 17154..17527
 /rpt_family="ERV1"
 repeat_region 18287..18333
 /rpt_family="MIR"
 repeat_region 18931..18972
 /rpt_family="(TTTA)n"
 repeat_region 18946..19246
 /rpt_family="Alu"
 repeat_region 21338..22611
 /rpt_family="L1"
 repeat_region 22593..22614
 /rpt_family="AT_rich"
 repeat_region 22814..23028
 /note="match to EST AA476419 (NID:g2204630) zx02b02.s1"
 repeat_region 22820..23263
 /note="match to EST N95327 (NID:g1267617) zb62e10.s1"
 repeat_region 22967..23324
 /note="match to EST N95327 (NID:g1267617) zb62e10.s1"
 repeat_region 23587..23944
 /note="similar to Bos taurus EST BE749870 (NID:g10163862)"
 repeat_region 24073..24534
 /note="match to EST H92027 (NID:g1087605) y682d10.r1"
 repeat_region 24352..24757
 /note="match to EST AA446796 (NID:g1259461) zw89f05.r1"
 repeat_region 24607..24757
 /note="match to EST AA476520 (NID:g2204731) zx02b02.r1"
 repeat_region 25062..25128
 /note="match to EST W07375 (NID:g1281376) za96d04.r1"
 repeat_region 25660..26105
 /rpt_family="GA-rich"
 repeat_region 26105..26147
 /rpt_family="L1"

```
repeat_region /rpe_family="(TA)n"
26712..26739
/rpe_family="AT_rich"
misc_feature 27181..27553
/note="match to EST W07375 (NID:g1281376) za96b04.r1"
misc_feature 27181..27199
/note="match to EST AA476520 (NID:g12204731) zx02b02.r1"
misc_feature 27465..27554
/note="match to EST BE296886 (NID:g9180453)"
misc_feature 27483..27554
/note="match to EST BG194829 (NID:g13716644)"
misc_feature 29343..29629
/note="match to EST BE296886 (NID:g9180453)"
misc_feature 29343..29564
/note="match to EST BG194829 (NID:g13716644)"
repeat_region /rpe_family="AT_rich"
31551..31764
/rpe_family="MIR"
repeat_region /rpe_family="MIR"
32002..32378
/rpe_family="MER2_type"
repeat_region /rpe_family="MER2_type"
32379..32588
/rpe_family="L1"
repeat_region 32589..32908
```

Query Match 45.3%; Score 1537; DB 9; Length 43877;
Beet Local Similarity 73.2%; Pred. No. 1.3e-261;
Matches 2554; Conservative 0; Mismatches 750; Indels 187; Gaps 38;

```
1 TGCTGCTCCGAGCATGAGAAAGTTGCAATCCCGGACGTGGAGAGAAACACGCTC 60
33032 TTCTATTTGGGGCTCAGAAAGTTGCAATCAGCCACGTGAGAAAGAAATTAACCTA 32973
61 TTGAATTTCTGAGTACCTCTTAATTCATTAAGTACCTCCGGA-----GATTGG 114
32972 CTTAAATTTCTTCAAGTTCTTCTTCAACCATTAAGTCTCTAGAGAAAGTTGATTAG 32913
115 ATAGAGCTGCTGCTATTTTCAAGTGTATGTTCTTCAACCATTAAGTGTGCAACA-TGA 173
32912 ATATAGTCATTTCTCCCTATTCATGATGTTATGTTCTTAAAGCATTTGCCAACATCGA 32853
174 ATAGATGTCCTGAGAACAGCTGCCCTTAAAGGAACTAAAGTTAGATTCCTCGACGCT 233
32852 GTTAGCAATATTTGAATGATTTGTTCCAGGGGAAAGAGGTTAGAGCTCTG--- 32796
234 CTGCTCAGTTTTTGTATACCAATCAATTAATTAACCTGTTTGTGTCATTTCTGTTT 293
32795 GTCAATACATTTTACATCCCAATCAATTAATTAACCTGTTTGTATGTTATGTTTGT 32736
294 AAAATATCTTATTTATAGTACTGCTAATTTCTTCAACATTTGGTTCAACCAAAAGG 353
32735 CAAATTA-CATTAATTTAATATATATGTTCAATTAACATTAACATCAAGCA---G 32681
354 CCAATTAATAGAGAGCCCTGAATGAGCTTACATTAACATTAACATTTGTTTCTATGAGA 413
32680 CAGACATTAATCAAGGCTTAATGATGTTATCTTACAGCA---TGCAATTTTTCATAG 32624
414 AAAATTTTCTCTGAGTCTGCAAGGCTTCTTGTAAATTTGAGACAAATATATACA 473
32623 AACTTTTTCCTTAGGCAATCAAGGCTT-----TTAAATGTTGTTAATATATACA 32569
474 TAAATGAACTGACTATTTAATTAATGATTTTAACTGTAAGTTCAAGTGCATTAAGTA 533
32568 ACATTTAGCA-----TCCTTAACATTTTATGAGTGTACAGTTGCGTGCATTAAGA 32518
534 CACTTACCTGCTGAGAACTATGATCAACATTAATCTCTAGACCTTTTATCTCTCG 593
32517 CATTCACAGTTGTGTAACATCAACATTAATCTCTAGAAATTTTATCTCTCCA 32458
594 AATTGAAACTCTGATCTATTTGACAGTAGCTTCCCATCAACAGCAAGCTTCTCTC 653
32457 GACTG-AACTCTGATCTATCAACAGTAAGTACCTCTCA-----GTGCTCTACC 32408
```

```
654 CAGCCCCCGGACCAACCATCTTCTTCCATGACAGTCACTCTGTGCTTAGAACA-- 710
32407 CAGTCCCTGGCAACACCATCTCTTCTTCCATGACAGTCTCTGTGATTTAGAACTAG 32348
711 -----CAGCCTTCAACACTACGCTTGGGGGCACTG---AAGCAACAGATC 756
32347 GAACACACGAGCAGCAGTCAAGCTATGCTTGGGGCCATTTTAAAAAGCAATTCAT 32288
757 ACTCCCTACCCGACCAAAATGACACAAATATATAAAAGCATGTGATATGATTCG 816
32287 AAGAGAGCATATAAAAAAGAGCAAAATATGTAAGAAATGCACTAATATGACGG 32228
817 AAAAAAGGCTGTTGCT--AAGTATGAGGCTGAAC--AAGCAAGAAATGACTAGT 871
32227 AAAAAAGGAATTTGTTATATGATATGAGAGCTGAAGAGGGAACCTCGCTTGT 32168
872 TGACCTCAGCTGGGATCCTGTGTG--AAGCTCAAAATTTTCAATGTTCTGTGAT 929
32167 TGACCTCAGCTGGGAACTGTGTGCTGATATGAGCTCAAAATTTTGGCCTCTGTGAT 32108
930 AAGCAAAATGCTTATTAAGCACTGTAAGATTTGATTAAGTAAATATCTAG 989
32107 GCCAGCAATGACCTGAAGAA---TGCTAGGAATATGATGTTGATTAATTTAGC 32052
990 CAAGCATTAATTTGCAAGCAGGATTCATGATTAACAGCACTGACAT--GTGAAA 1047
32051 AAGTAGCAATTTGTGAATACAGAAATCATGAATTAAGAAATGTTCTGTAGCGAAA 31992
1048 TGATATCTTGTGTTCTTATTTCCAGGAGTAAGGAAAGGCG-----T 1093
31991 TGACAACTTTGTCTCTTTTTCAGGTATTAAGAAATTAAGACCTGAAACATCTTGT 31932
1094 CACAGGCTGCTTACACATTTTCTAGAGAGCTAGCTATGTCAGTGG----- 1144
31931 AACCTGGCTGCTTGAAGAAATTTTACTGCTCACAAGCTTAAGAAAGTATGATGAC 31872
1145 ---TAGCTGCAATTAACAAGTGAAGATTTCTAGTGTGAGAGTGAATTTAACCA 1201
31871 TGATTAATTTGAGCTATGAGCTGAAGAGTCTGATGATGATGAGATTAATTTAAGA 31812
1202 TAACTCAACCCCTCTGCAATGAAGAGCTAAGTCAATGATACAGTATCAAGCA 1261
31811 TAACTCAACCCCTCTGCAATGAAGAGCTAAGTCAATGATACAGTATCAATTTA 31752
1262 CTTACTATATGACAGGCACTGTAACAATTTACATGATTAATTAAGTATTAACAGCCC 1321
31751 CTTACTATATGACAGGCACTGTAACAATTTACATGATTAATTAATTAATTAATTAATTA 31692
1322 CAACACTCTATGAGAGAGCTGAAGTTAG--AAGTATCTATCATTTATTAACA 1376
31691 CAAAACCTATGAGAGAGCTTAAGTTAGGAGTAAATGATTCATTCATTAACATTA 31632
1377 GT-----GGCAATGAGATGTAAGTCAAGTCAATCAATCCAGAGCTGAGATC 1428
31631 GTTAGAAGTGCAGAAAGTTGAGATTTGAATCTAGGTTATCTGACTCCAGAGCTGAGTTC 31572
1429 CCAATTCACACAAATCTATCAAGTAAAGGAAAGAAAGATTTGATTTGCTCAGAG 1488
31571 TCAATTCACAGTATCAATTTCAAGCATATTAAGAAAGAAAGTTGATCACTTGTGAC 31512
1489 TGTATAGGGCATATGTTACA--ATTATTAACATTAACAAGATTTATATGTTAAAAATA 1545
31511 TGTATATGATGATGTTTAAACAATGATTAACATTTGAAAGATTTTACATTTGAAATATG 31452
1546 AATTATCAAAATTAACCTTATTAAGCTGATCAATATGCTCCGCAACAAAGACT 1605
31451 AATTATCAAAATTAACCTTATTAAGCTGATCAATATGCTCCGCAACAAAGACT 31392
1606 ATCTGAATCTCTCAGGGCATGTTGTTGTTGTTCTTAACTTTAAATGATGG 1665
31391 ATCTGAATCTCTCAGGGCATGTTGTTGTTGTTCTTAACTTTAAATGATGG 31333
1666 CAAATTAATGATTAATGTAAGGCA-TTTTTTCTCAAGAGATGATATGATCTCTTAAGA 1724
```


Db 126911 TTGAAAAAGCTAGAAATATTGAGTTTCATGATTAATTTCAGC---AAGAGGCAAT 126856
Qy 1002 GTGCAGACGGAGATCCATGATTAACAGACCTGACCAT--GTGAAATGATTAATCTTTG 1059
Db 126855 TTGTGATATCAGAAATCATGATTAATGAAATTGTCTGTAAAGTGAATTCACATCTTTG 126796
Qy 1060 TTTCCTTTATTCAGGAGTAAAGGAAAGCGC-----TCACAGGGCTGCC 1105
Db 126795 TTCACTTTTTCCAGGTAGTAAGATGAAGAGCTGGAACATGCTTTGTAACCTGGCTGCC 126736
Qy 1106 TTACACCACTTTTA-----CTAGAGAGCTGACCTATGTCACTGGTAGCTGG 1151
Db 126735 TTAGAAATTTTATCTTGCTCCACAGACCTGAAGAGTGAACCTTGTGACT--GATTAATGG 126677
Qy 1152 CAATTACAACTGAAGAGAGTCTAGTTCATGTGAGATGAATTAACATTAATCTTAAC 1211
Db 126676 CAGCTATATCTTGAGAGAGTCTAGTTCATGTGAGATGAATTTTTCAGATTAATCTTAAC 126617
Qy 1212 CCCCTCTGATGAACAGAGACTAAGTACTAAGTACCAAGTATCACTTAATAT 1271
Db 126616 TCCTTCTGATTAACAAAGAGAGCAAGCACTCAAGTCCAGTTATCAATTAATCTTAATAT 126557
Qy 1272 GACAGGACTGTATCAACAATTTATGATGATTTATGATTAATCATGCCCCCACTCTA 1331
Db 126556 GACAGGTGTATCATTCAGCAATTTATCATGATTTAATTAATATCCCCCAAAACCTTA 126497
Qy 1332 TGAGAGAGCTGAAGGTAGAG----AAGTATCTCAATTCATTAATACAGT----- 1378
Db 126496 TGAGAGAGCTGAAGGTAGAGGTGAAGTGAAGTATCTCATATCATTAATTAAGTGAAGTG 126437
Qy 1379 GGAACAATGAGATCTGAACCTGAGTCTATCCAACTCCAGACCTGAGATCCCAAT----- 1433
Db 126436 GCAAGTGTGAATTTGAACCTGAGTCTATCTGACTCCAGAGCTGAGTTTCAATTCAC 126377
Qy 1434 TGTCTACAAATTTCTATCATAGTTAAAGGAAAGGATTTGATTTGCTCAGAACTGAT 1493
Db 126376 TGTCTATGAATTTCTAAGCATTTAAAAAAGTTGATTTGATTTGCTCAGAACTGAT 126317
Qy 1494 AGGGGATATGTTACA---ATTATTAACATTAAGGAAAGTATTAATGTTGAAATAATTT 1550
Db 126316 AGATGCAATGTTTACATGATCATTAATTTCAAGAGATTTTACACATTAAGAAATTAATTT 126257
Qy 1551 ATCAAAACAATTAACCTTTATTAAGCTGATCTAATATCTCTCCGCAACAAAGACTATCTG 1610
Db 126256 ACCAAACAATTAACCTTTAAGCCAGATCTAATCTGATCCCAATACAAAGAGATCTG 126197
Qy 1611 AAATCTTACAGGCACTGCTGTTGTCTGCTGTTTCTTAATCTTTAATGATGGCAAT 1670
Db 126196 AAATCC-TCAGGGCACTGCTGTTGTCTGCTGTTTCTTAATCTTTAATGATGGCAAT 126138
Qy 1671 CTAATGCAATTAATTAAGGCA-TTTTTCTCAAGAGATGATACCTCTTAAGAAATTTG 1729
Db 126137 CCAATGCAATTAATTAAGGCAATTTTTTCTCAAGAGATGATACCTCTTAAGAAATTTG 126078
Qy 1730 ATGAAATGATTAATCTTTACAGGCTACTGATGATTTTATGTCATGAGGCAATTA 1789
Db 126077 ATGAAATGATTAATCTTTACAGGCTACTGATGATTTTATGTCATGAGGCAATTA 126018
Qy 1790 TTAGTGTCAATGTCGCAAA--GTAGTCACTAAATAATTAATTTGATGATGACAC 1846
Db 126017 TTAGTGTCAAGTGTGCAAAATGTAGTCACTTTAAATAATTAATTTGATGATGACAC 125958
Qy 1847 TGCACTCTCTGGGGGAAAAAGTATGATTAATCTCTTAAGGATCCCTTAAGCTCCCA 1906
Db 125957 TGTATCTCTTGGGAAAAAGTATGATTAATCTCTTAAGGATCCCTTAAGCTCCCA 125898
Qy 1907 AAAGTAGAGAAAAATTAATCTCTGCTGCGCTGAGAAACAGCTTCTGTTTCTGTGCT 1966
Db 125897 AAAGTAGAGAAAAATTAATCTCTTGTGCTGAGAAATCTCTGTTTCTGTGCTG 125838
Qy 1967 ATATTGTTTAGTTTAAATGTTTCAATTTGATTTGACCTTGTGCTCCCAAGTAAAG 2026
Db 125837 ATGTTGCTTAGTTTCTTAATAGTTCATTTGATCTGTGTGCTCCCAAGTAAAG 125778

Qy 2027 TTGAGAGTTGATCCCTACAGAGGCCACTTCAATTTAAGAAACAAAAAGCCCATCTCT 2086
Db 125777 TTGAACGTTTATCCCTACAGAGGCCACTTCAATTTAAGAAACAAAAAGCCCATCTCTCT 125718
Qy 2087 GCTCCAGACCTTAACCCCAATTCCTGCGAGGTGTCTGCTCCGCTCAATGAGAACT 2146
Db 125717 GCTCCAGACCTTAACCCCAATTCCTGCGAGGTGTCTGCTCCGCTCAATGAGAACT 125658
Qy 2147 GGCAGAGAACTACTAGAGAGGTGCGACAGTACTGAAAGTGAATAATGACTGACAC 2206
Db 125657 TGGCAAGAGGTGAACA-----TATCAGAG 125631
Qy 2207 TACTGAGAGGAGAAAAATGAGGACCCCTCATGATGATGTTCTTTCCCTTCTGTCTC 2266
Db 125630 TATTGGAGAAACAAAGAGTCAACCCCTTTATCATGATGCTTTCTCTTTAGCTGCTC 125571
Qy 2267 ACAATGCTCGATTAATTTAACAAGGTAGATA-----ACTCAATTTTTTCTTTT 2318
Db 125570 ACAATATCTGATTAATTTAAGAGAGATGATCTGCACTTTTAACTCTCGAATAATTT 125511
Qy 2319 ACCACTGAGAGCTGAGAGAACTTTGTTACCCATCATTAATTAATCATATCTTCAAGTC 2378
Db 125510 GCTGCTAAGAACTGAGGCAACCTGTAATTAAGCTCTGCAATGCAATTAAGTCAAT 125451
Qy 2379 ATTCTATGTTATTTCAAGA-----TCAATAGCTGACATATCTCTTTGTAATTAACA 2432
Db 125450 ATCTTATTTCTATGTTCAAGATTTTTCATAGTCAATATATCTCTTGTGTAATTAACA 125391
Qy 2433 ATGAAAAACACATCTCTGAGCAATTAATCTGCAACTTTAGATGAGAACTTA 2492
Db 125390 ATGAAAGAAACACATCTCTGAGTAAATTAATCTGCACTTTAGAAAGTGA--ATA 125336
Qy 2493 ATACTAGCATTAAGAACTGAATAACAATTTCAATGAATAAAGATAT-----TATTT 2547
Db 125335 ATAGACTCACTCAAAAGTCAAGTAAATTTTCATATTAATTAAGCAATTAAGTAACTATGT 125276
Qy 2548 AAAAGTATTCATAGAGCAATTTAATTA--AAGTAAAGTTTCTATTAATGATTAAGAT 2606
Db 125275 AAAATATCTTCATAGCAATTAATGATTAATTAAGTAACTTAATTTACTAAGAT 125216
Qy 2607 TTAATCAGGAAACAAGTTTCTCAATTAATTAAGCAAAAACTTTTACTAGTACAGTC 2666
Db 125215 TTAATCA-AAAAACAAGTTTCTCAATTAATTAAGCAAAAACTTTACTAGTACATTAATTC 125157
Qy 2667 TTTTCATTAAGTC-TTCCGTAATTAATCTGATTTTCTAATTAATTAAGCAAGTAAATTA 2725
Db 125156 TTAATTTTAAATTAAGTCTTCTTAATTTTCTAATTAATTAAGTAACTTAAGTAAATTA 125097
Qy 2726 ATTTAATTAACAATAATTAATTTTACTCAAAATCTTAATTAATTAAGTAAATTAATC 2785
Db 125096 ATTTAATTAACAATAATTAATTTTACTCAAAATCTTCTTAACAATAATTAATTAATC 125037
Qy 2786 ATTTAATTTCTGAGGAAAAAGCATATCACTTTTAAATTAAGTAAATTAAGTAAATTA 2845
Db 125036 ATTTAATTTTGAAGAAATTAATTTTCTCAAAATCTTCTTAACAATAATTAATTAATTT 124978
Qy 2846 TATTCATTAATTAATTTTAAAGTTTCAATTAATTAAGTAAATTAAGTAAATTAAGTAAAT 2903
Db 124977 GACTTATTTAATTAATTAAGTAAATTTTAAAGTAAATTAAGTAAATTAAGTAAATTA 124928
Qy 2904 -GTATATGTTATGTTAATTAAGTAAATTTTAAAGTAAATTTTAAAGTAAATTAAGTAAAT 2962
Db 124927 AGTTTATTTATTTGCAACATTAAGTAAATTTTAAAGTAAATTTTAAAGTAAATTTAAT 124868
Qy 2963 TTCAATGATTTATTTCTTTTAAAGATTAAGTAAATTTAAAGTAAATTTAAAGTAAATTT 3022
Db 124867 TTTGAGATTTATTTCTTTTAAAGATTAAGTAAATTTAAAGTAAATTTAAAGTAAATTTAAT 124808
Qy 3023 AAAATATCTGTTGAGAACTTGTGACAGAGGTTTAACTCTGACAGGAGATTTCAAT 3082
Db 124807 AAAATATCTGTTGAGAACTTGTGACAGAGGTTTAACTCTGACAGGAGATTTCAAT 124748

```

OY 3083 GTGAGCAGAGCCCATCATCAGATCCCGACGACCTTGTCTCATCAAAAGTTGGAATATA 3142
DB 124747 GTGGAGCAGAGCCCATCATCATGATCTGACGACCTGTCTCATCAAGTTGGAATATA 124688
OY 3143 AAGGCACTTGGATATACATATTAAGATTCTGCTGGCAAGTTGCTCTGACGCTG 3202
DB 124687 AAGGCACTTGGATATACATATTAAGATTCTGCTGGCAAGTTGCTCTGACGCTG 124628
OY 3203 GGCAGGCACTTAACGTTTGGCTTGGCTTACTCMAAGCAAAAGAAAGTAAGAGAGNA 3262
DB 124627 TACATGCACTTAATATTTTGTCTGGCATTACTCAAAAGCAAAAGAAAGTAAGAGAGNA 124568
OY 3263 GTAAGAACAG-GGAAAAGATTGTATGTATTTAAACATCCAAAAGTCTGATCTCT 3321
DB 124567 ATAAAGAACAGAAAAAGATTATATGTATTTAAATCATGCCAAAAGTCTGATCTCT 124508
OY 3322 GTTATATTTTACCTATTTATGCTGATTTGCTGGCCGAGTGATCTGATGAGAACAGC 3381
DB 124507 GTTATATTTTACCTATTTATGCTGATTTGCTGGCCGAGTGATCTGATGAGAACAGT 124448
OY 3382 GAGCAGAA 3389
DB 124447 GAGCAGAA 124440

```

```

RESULT 10
LOCUS AC142556 153029 bp DNA linear HTG 04-APR-2003
DEFINITION Pan troglodytes clone CH251-7606, WORKING DRAFT SEQUENCE.
ACCESSION AC142556
VERSION AC142556.1 GI:29540591
KEYWORDS HTG; PHASE2; HTGS; DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

```

```

REFERENCE
AUTHORS 1 (bases 1 to 153029)
Ahter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Carlaea,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B.,
Marquis,E.H., Masello,C., Maskeri,B., McDowell,J.,
Pasulnigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stentrop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wehrby,K.D., Wiggins,L., Young,A. and Green,E.D.
TITLE NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 153029)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2003) NIH Intramural Sequencing Center, 8717
Grosvont Circle, Gaithersburg, MD 20877, USA
COMMENT ----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: dhg
Center clone name: 076006
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 153025 bases at least Q40
Consensus quality: 153026 bases at least Q40
Insert size: 123000; agarose-fp
Insert size: 153029; sum-of-config
Quality coverage: 15.91x in Q20 bases; agarose-fp
Quality coverage: 12.79x in Q20 bases; sum-of-configs

```

```

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 153029: contig of 153029 bp in length.
FEATURES
source
1..153029
location/Qualifiers
1..153029
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-7606"
/clone_11b="CH251"
1..153029
/clone="assembly_fragment"
/clone_end="SP6"
vector_side="right"

```

```

ORIGIN
Query Match 44.8%; Score 1519.6; DB 2; Length 153029;
Best Local Similarity 72.9%; Pred. No. 1.3e-258;
Matches 2540; Conservative 0; Mismatches 764; Indels 182; Gaps 38;
OY 1 TGCCTCCGAGGACATGAGAGGTTGGCAATCCCGGACGAGGAGGAAACACGCTC 60
DB 42782 TTCTATTTTGGGCGATCAGAGAGTTGGCAATCACTCCGCTGAGAGAAATTAACCTTA 42841
OY 61 TTGAATTTCTTGAGTACCTCTTAATTCATTAATGCTGACCTCCGAG-----GATTGG 114
DB 42842 CTTAAATTTCTTGAGTACCTCTTAATTCATTAATGCTGACCTCCGAG-----GATTGG 42901
OY 115 ATAGAGCTGATCTCATATTTATTCAGAGGTTATGTTTCAACCAATCACTGCAACA-TGA 173
DB 42902 ATATAGCTGATCTCATATTTATTCAGAGGTTATGTTTCAACCAATCACTGCAACA-TGA 42961
OY 174 ATAGAGCTGATCTCATATTTATTCAGAGGTTATGTTTCAACCAATCACTGCAACA-TGA 223
DB 42962 GTTACGAAATATTTGAATCTATTTGTTCCAGGAGAAACAGGTTATGTTAGCCTCTG--- 43018
OY 234 CTGCTCAGCTTTTGTATTAACCAATCAATTAATTAATGTTTGTGTCATTTCTGTTT 293
DB 43019 GTCAATACATTTATTAATCAACCAATCAATTAATTAATGTTTGTGTCATTTCTGTTT 43078
OY 294 AAAATATCTTTATTTAATGCTGATCTTAATTTCTTCAACATTTGTTTCAACCAACAAAG 353
DB 43079 CAAATTA-CATTAATTTAATTTAATTTGTTTCAATTAATGCTGATCTTCAACCAACCAAGC 43137
OY 354 CCCATTAATCAAGAGCCCTGAATGAGCTTACATTAACCAACATGTTTCTTATGAGGA 413
DB 43138 ACC-----ATACTCAGGCTTGAATGAGCTTATCTTATGATTA--TGCAATTTTCTATAG 43190
OY 414 AAAATTTTCTTCTGATCTTCCACAGGCTTCTTGAATTTGTTGACAAATATATACA 473
DB 43191 AACTTTTCTTCTGATCTTCCACAGGCTTCTTGAATTTGTTGACAAATATATACA 43245
OY 474 TAAATGAACTGATCTTATTTAATGATTTTAATGATTTTAATGATTTTAATGATTTAATGTA 533
DB 43246 ACATTTTGA-----TCTTAACATTTTATTTGTTTGAATGATTTGATTTGAATGA 43296
OY 534 CACTTACCTGCTGAGAACTATGATATCAATGATCTTATGATCTTATGATCTTATGATCTTATG 593
DB 43297 CATTCACCTGTTGATATCAATCAACCAATCTTATGATCTTATGATCTTATGATCTTATGATCT 43356
OY 594 AATTGAAACTCTGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATG 653
DB 43357 GACTG-AACTCTGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATG 43406
OY 654 CAGCCCCCGGACCAACCATCTTCTTTCATGACATCTCTGCTGCTAGGAACA--- 710

```



```
Db 43407 CAGTCCCTGGACACCACTGACTTTCATTTGACGCTTTCCTGTGATTAGAACTATG 43466
Qy 711 -----CAGCCCTTCAACACTAGCGCTGGGGGGACGTG----AAGCAACGATC 756
Db 43467 GAACACACGAGCAGTGTAGACTATGCTGGGGGCAATTTTAAAAAGCAAAATCAAT 43526
Qy 757 ACTCCCTACCGCCACCAATGCAACAAAAATATTAAGACATGGTGCAATTCATGTC 816
Db 43527 AAGGAGCAATTAATAAAGAAAGCAAAATATGTGAAAACATGGCACTTAATATAGTGG 43586
Qy 817 AAAAAAGGTGCTGCT--AAGTATGAGGCTG--AAACAAGCAGAGAAATGACTAGT 871
Db 43587 AAAAAAGGAATTTGTTATAGATGAGAGCTGAAAAAATTAAGCGGAACCCGCTTGT 43646
Qy 872 TGACCTGAGCTGGGATCCTGTGTGTGG--AAGCCTCAATTTTCATTTGCTGTGTGAT 929
Db 43647 TGACCTGCTGGGAACCTGTGCTGATGAGATTCAAATTTTGTGCTGTGTGAT 43706
Qy 930 ACGCACAATGCTTATTAAGCACTGTAAAGATTGATTATGAAATTAAGATTAATCTG 989
Db 43707 GCCCGCATGACTGAAAAA-----TTCTAGGAATTTGAGTTTCATAGTAATTTTAC 43762
Qy 990 CAAGACATTAATGTGCAAGCAGCGGATCCATGATTAACGAGCACTGACCAT--GTGAAA 1047
Db 43763 AAGTAGGCAAAATTTGTGATACAGAAATCATGAAATTAAGAAATGTCTGTAGGGGAA 43822
Qy 1048 TGATTAATTTTGTTCCTTTATTCAGGCAATTAAGAAAGGCTC-----1094
Db 43823 TGACAACTTTTGTTCATTAATTTTCCAGGTAAATTAAGAAAGCTGGAACATGCTGT 43882
Qy 1095 -ACAGGGCTGCTTACACATTTTACTAGAGCTAGGCTATGTGCTGCG-----1144
Db 43883 AACTGGCTGCTTGAATAATTTTACTTGTCCACAGGCTTAGAAGTAGGCTTGTGAC 43942
Qy 1145 --TAGCTGCAATTAACAACCTGAAGAGTTCTAGTTCAATGTGAGATGAATTTTAA 1201
Db 43943 TGATTAATTTGAGATGACCTGAAGAGTTCTAGTTCAATGTGAGATGAATTTTAA 44002
Qy 1202 TAAATTCACCCCTCTGATGAACAGAGCTAGTACTAGTACCAAGTCAATTAAGTCA 1261
Db 44003 TAAATTCACCCCTCTGATGAACAGAGCAAGCACTCAAAAGTCCAGTTATCAATTA 44062
Qy 1262 CTACTATATGACAGGCACTGACTCAACAATTTTCAATGATTTATGAATTCATGCC 1321
Db 44063 CTACTATATGACAGGCTGCAATTTTCAAGCAATTTTCAATGATTTATTAATTC 44122
Qy 1322 CAACACTATGAGAGCTGAAGGTTAG--AAGTATCTCAATTCATTTATTAACA 1376
Db 44123 CAAAACCTATGAGAGCTGAAGGTTAGGGAAGTTAGTATCTCATCATTAACAATA 44182
Qy 1377 GT-----GGCAACCTGAGATCTGAAGTCTGATCTCAATCTCCAGAGCTGAGATC 1428
Db 44183 GTTAAAGTGGCAAGGTTGAGATTTGAACCTCAGGCTTCTGACCTCCAGAGCTGAAGTTC 44242
Qy 1429 CCAATTTGATCACAATTTCTAATCAAGTTAAAGGAAAAAGATTGATTTGCTCAGAG 1488
Db 44243 TCAATTTCACTGCTATGCAATTTCTAAGCATTTTAAAAAAGTTTGACTTACTGGAAC 44302
Qy 1489 TGTATAGGGGCAATGTTACA---ATTATTAATTAACAAGATTATATGTTGAAAAATA 1545
Db 44303 TGTATAGATGATGTGTTACATGATCATTAATTTCAAGATTTTACATGAAAAATG 44362
Qy 1546 AATTATTAACAAGATTAACCTTTATTAAGCTGATCTAATCTGCTCCGCAACAAGACT 1605
Db 44363 AATTATTAACAAGATTAACCTTTATTAAGCAAGATCTAATTTGCTCCATTAACAAGAT 44422
Qy 1606 ATCTGAATCTTTCAGGGGCACTGGTTGTGTGCTGTTTCTTATCTTATATGATGG 1665
Db 44423 ATCTGAATCTTTCAGGGGCACTGGTTGTGTGCTGTTTCTTATCTTATATGATGG 44481
Qy 1666 CAATCTATGCAATTAATGTAAGCA--TTTTTCTCAAGAGATGATGATCTCTTAAGA 1724
|||||
```

```
Db 44482 CAATTCATGATTAATGTAAGCAATTTTCTCAAGAGATGATGATACCTTTAAGA 44541
Qy 1725 AATTGATGAATGATTAATTAATCTTTTCAAGGCTACTGATGATTTAGTGAAGCA 1784
Db 44542 AATTGATGAATGATTAATTAATCTTTTCAAGGCTACTGATGATTTAGTGAAGCA 44601
Qy 1785 GTAAATTAATGATTAATGATGCAAA--GTAGTGAATTAATTAATTAATTAATGATGA 1841
Db 44602 GTAAATTAATGATTAATGATGCAAAATGATGATGATTTAAATAATTAATTAATGATGA 44661
Qy 1842 ACCACTGCACTCTCTTGG--GGAAAAAAGTATGATTAATCTCTTGAAGATCTTACT 1900
Db 44662 GCCACTGATTTCTCTTGGAAAAAAGTATGATTAATCTCTTGAAGATCTTACT 44721
Qy 1901 TCCCAAAAGTATGAAAAAATTAATCTCTGATGCTGGAACAGCTCTGTTCTTG 1960
Db 44722 TCCCAAAAGATGAAAAAAGTATCTCTTGGCTGGAATATCTTCTGTTCTTG 44781
Qy 1961 CTGGCTATTAATTTGTTAGGTTTAAATAGTTCAATTTGATTAAGCTTGTGCTCCCAAG 2020
Db 44782 CTGGCTATTTGTTAGCTTACTTTAAATGTTCAATTTGATTAAGCTTGTGCTCCCAAG 44841
Qy 2021 CTAAAGTGAAGTTGATCCCTACAGAGGCACTTCAATTAAGAAACAAAACCCCA 2080
Db 44842 CTAAAGTTGAACGTTTGAATCCCTACAGAGGCACTTAAATTTAGAAACAAAACCTCA 44901
Qy 2081 TTCTGTGCTCCAGCCTTACCCCAATCCGTCAGAGGTCTGCGCTCCGTCAAATGA 2140
Db 44902 TTCTGTGCTCCAGCCTTACCCCAATCCGTCAGAGGTCTGCGCTCCGTCAAATGA 44961
Qy 2141 GAAACTGCAAAAGAAAGTACTAGAGGTGCGACAGTACTAGAAAGTAAATGAGACTA 2200
Db 44962 AGAAATTTGCAAAAGGAGTGAACATATTGAG-----44994
Qy 2201 GCACTACTAGAAAGCAAAAAATGAGCACTTCAATGATGATCTTCTTCTCT 2260
Db 44995 -----TATTTGGAACCAACAAAGTCAACCCCTTATCAATGATGCTCTTCTCTTAT 45048
Qy 2261 GTGTTCACATGCTCCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2312
Db 45049 GTGTTCACATTAATTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 45108
Qy 2313 TCTTTTACACCTGAAAGCTGAGAAAATTTGTTTACCATCATTAATTAATTAATTA 2372
Db 45109 ATATTTTACCTGCGGAATCTGAGCAACCTGAATTAATTAATTAATTAATTAATTA 45168
Qy 2373 TAACTATTTATGTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2432
Db 45169 TTCAATTAATTTATGTTTACAGAAATTTTCTACTAGCTGACATTAATTAATTAATTA 45228
Qy 2433 ATGAAAAAACAATCTCTGAGCAATTAATTAATTAATTAATTAATTAATTAATTA 2492
Db 45229 ATG--AAAAACATTTTCTGAGTTATGTTAATTTGATTTGATTTGAAATGAA--ATA 45282
Qy 2493 ATACTAGTCAATTTGAACCTGAATTAATTAATTAATTAATTAATTAATTAATTA 2547
Db 45283 ATGACACTCAATCAAAAGTTGATTAATTTTCAATTAATTAATTAATTAATTAATTA 45342
Qy 2548 AAAAGTAAATTCATGAGCAATTAATTAATTA--AAGTGAATTTTCTTAATGCTTAAGAT 2606
Db 45343 AAAAATAATTCATGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 45402
Qy 2607 TTATTCAGGGAACAAAGTTTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTA 2666
Db 45403 TTATTCAGGGAACAAAGTTTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTA 45461
Qy 2667 TTTTCATTTTAAGTC--TTCTGAATTAATCTGATTTTCTAAATTAATTAATTAATTA 2725
Db 45462 TTAATTTTAATTAATGAGCTTCTGATTTTATTTTCTTAATTAATTAATTAATTAATTA 45521
Qy 2726 ATTTAATTAATTAATTAATTAATTTT--TACTTCAATGCTTAAATTAATTAATTAATTA 2784
Db 45522 ATTTAATTAATTAATTAATTAATTTTCTACTTCAAAATTAATTAATTAATTAATTAATTA 45581
|||||
```


Db 140181 TTCTATCTGGGGGATCAGGAAGGTGGCAATCACTGCACCTGGAGAAAGAAATTA 140122
 Oy 61 TTGAATTTCTGAGTACTCTTAAATGCAATGCACTGCGA-----GATGG 114
 Db 140121 CTAAATTTCTGAGTCTTCTTCCACCAATCAATGCTTCTAGGAAGTTATG 140062
 Oy 115 ATAGAGTCACTCATATTAATCAAGTGTATGTTCTACCAATCACTGCGACA-TGA 173
 Db 140061 ATATAGTCAATTTCTCCATTCATAGTAGTATGTTCTATTAAGTCAATGCGACAC 140002
 Oy 174 ATAAAGTATCTGGAACAATGCCCCCTAGGGGAATCAAGATTGATTCCTGACCT 233
 Db 140001 GTTAGCAATATGGAATGATGTTCCAGGGGAAAAACAGGGTATGAGGCTGTG--- 139945
 Oy 234 CTGGGACGTTTTGTTAAACCAATTAATTAACCTGTTTGTGTGATTCGTTT 293
 Db 139944 GTCAATACATTTAATCAACCAATCAATTAATTAACCTGTTTGTATGTTATTTA 139885
 Oy 294 AAAATATCTTATTTAATAGTACTGTAATCTTCAACATTTGGTGCACAACCAAG 353
 Db 139884 CAATTA-CATTAATTTAATTAATTTGTTCAATTAACAATGAACTCAAGCGCAACG 139826
 Oy 354 CCTAATTAATGGAAGCCCTGAATGAAGCTTAATTAACAACATTTGTTTCTATGAGA 413
 Db 139825 ACC-----ATACTGAGCCTGAATGATGTTATCTAGCATTA---TGCAATTTTTCAT 139773
 Oy 414 AAAATTTTCTGCTGCTGCGACAGGCTTCTGCTTAAATTTGTGACAAAATTAACA 473
 Db 139772 AACTTTTTCCTTAGGATATCAAGGCTT-----TTAAATTTGTGTAATTAACA 139718
 Oy 474 TAAATGAATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 533
 Db 139717 ACATTTACA-----TCTTAACATTTTATAGTGTACAGTTGCGGTGATTAACA 139667
 Oy 534 CACTTACCTGCTGTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTA 593
 Db 139666 CATTCACACTGTTGTATTAACATCAACACATTAATTAATTAATTAATTAATTA 139607
 Oy 594 AATTGAATCTGTGATCTATTTGACAGTAGCTTCCCCCAATCAACCGCAACCTTCTC 653
 Db 139606 GACTG-AAACTCTGTATCTATCAACAGTAATCACTCCCTCA-----GTGCTCAAC 139557
 Oy 654 CAGCCCCCGGCAACCACTCTTCTTCAATGACATCTGCTGCTAGGAACA--- 710
 Db 139556 CAGTCTCTGGCAACCACTGCTACTTCTTCAATGACCTTCTCTGTATGGAACAATG 139497
 Oy 711 -----CAGCCTTCAACATGCTGCGGGGCACTGT---AAGCAACAGATC 756
 Db 139496 GAAACACAGGAGCACTGACACTATGCTGGGGCAATTTTAAAGCAAAATCAAT 139437
 Oy 757 ACTCCCTACCGGCAACCAATGACACAAATTAATTAATTAATTAATTAATTAATTA 816
 Db 139436 AAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 139377
 Oy 817 AAAAAGGAGGCTGCT--AAGTATAGGGCTG---AAACAAGGAGAAATGACTAGT 871
 Db 139376 AAAAAGGAGAAATTTGTTTATAGTATAGAGCTGAAAAAATTAAGCGCAACCTGCTTGT 139317
 Oy 872 TGACCTGAGCTGGGATCTGTGTGTG--AAGCTCAAAATTTTCAATTTCTGTGAT 929
 Db 139316 TGACCTGCTGGGAACCTGTGCTCAATAGATTAATTTTGTGCTGTGTGAT 139257
 Oy 930 ACGCAAAATGCTTATTAAGCACTGTAAAGATTGATTAAGATTAATTAATTAATTA 989
 Db 139256 GCCCGCAATGACTAAAAA---TTCTAGGAATTTAGTTTCAATGAATTAATTTAGC 139201
 Oy 990 CAAGCATTAATGTCAGACAGGATCATGATTAATGAAGCACTGACAT--GTGAAA 1047
 Db 139200 AAGTAGCAATTTTGTGATACAGAAATCATTAATTAAGAAATGTCTGTAGCGGAA 139141
 Oy 1048 TGATATCTTTGTTTCTTTATTCAGGAGTAAGAGAAAGGCTC----- 1094

Db 139140 TGACAACCTTTGTTCACTATTTTCCAGGTAAATGAATTAAGAGCTGAACTGTGT 139081
 Oy 1095 -ACAGGCTGCTTACACATTTTACTAGAGAGTACCTATGTCACTG----- 1144
 Db 139080 AACCTGCTGCTTGAATAATTTTACTTGTCTCACAGGCTTGAAGAGGCTTTGTAC 139021
 Oy 1145 ---TAGCTGGCAATTAACAATGAAGAGTTCTAGTTCAATGTGAGAGATGAATTTA 1201
 Db 139020 TGATTAATGGAGCTATGACCTGAAGCACTTCAATGATCAATGTGAGATTAATTTA 138961
 Oy 1202 TAATCTCAACCCCTCTGCTAGAAACAGAGCTAATGATCTCAATCACTTACATCA 1261
 Db 138960 TAATCTCAACCCCTCTGCTAGAAACAGAGCACTCAATGATCAATGATCAATTA 138901
 Oy 1262 CTATCTATGAGAGAGAGCTGATCAACATTTATCAATGATTAATTAATTAATTAATTA 1321
 Db 138900 CTATCTATGAGAGAGAGCTGATCAACATTTATCAATGATTAATTAATTAATTAATTA 138841
 Oy 1322 CAACACTCTATGAGAGAGCTGAGAGTTAG--AAGTATCTCATTTATTAATTAACA 1376
 Db 138840 CAACACCTATGAGAGAGCTAAGTTTAGGAGATTAAGTATCTCATCTATTAACA 138781
 Oy 1377 GT-----GGCAACTGAGATCTGAATCTAGGCTTATCCACTCCAGAGCTGAGATC 1428
 Db 138780 GTTAGAGTGGCAAAAGTTGAGATTTGAATCTAGGCTTATCTGAGCTCCAGAGCTGAG 138721
 Oy 1429 CCAATTTGCTACAAATCTATCAAGTTAAAGGAAAGGAAAGTTGATTTGCTCAGAA 1488
 Db 138720 TCAATTTCACTGCTATGCAATTTCTAGCATTTTAAAGAAAGTTGATTTACTTGAAC 138661
 Oy 1489 TGATAGGGGCAATTTGTTACA---ATTATTAATTAACAAGATTTATATGTTAAAA 1545
 Db 138660 TGATATAGTATGATGTTACATGATCAATCAATTTCAAAAGATTTACATTTGAAAA 138601
 Oy 1546 AATTTATCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1605
 Db 138600 AATTTATCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 138541
 Oy 1606 AATGAAATCTTCAAGGAGCTGATTTGATGCTGATTTGCTTAAATTTAATTAATTA 1665
 Db 138540 AATGAAATCTTCAAGGAGCTGATTTGATGCTGATTTGCTTAAATTTAATTAATTA 138482
 Oy 1666 CAAATCTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1724
 Db 138481 CAAATCTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 138422
 Oy 1725 AATTGATGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1784
 Db 138421 AATTGATGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 138362
 Oy 1785 GTAAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1841
 Db 138361 GTAAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 138302
 Oy 1842 ACCACTGACCTCTTGG--GGAAAAAGTAATTAATTAATTAATTAATTAATTAATTA 1900
 Db 138301 GCACTGATTTCTCTTGAATAAATAATTAATTAATTAATTAATTAATTAATTAATTA 138242
 Oy 1901 TCCCAAAAGTAGTAAGAAAAATTAATCTCTGTGGCTGGAACAGCTTCTGTTTCTTG 1960
 Db 138241 TCCCAAAAGTAGTAAGAAAAATTAATCTCTGTGGCTGGAACAGCTTCTGTTTCTTG 138182
 Oy 1961 CTGGCTAATTTGTTAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2020
 Db 138181 CTGGCTAATTTGTTAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 138122
 Oy 2021 CTAAAGTTGAGAGTTGATCCCTACAGAGGCACTTCAATTTAGAGAACAAAAAGCCCA 2080
 Db 138121 CTAAAGTTGAGAGTTGATCCCTACAGAGGCACTTCAATTTAGAGAACAAAAAGCTCA 138062
 Oy 2081 TTCTGTCTCCAGACTTAATCCCAATCTCTGCAAGGTGTCTGCTCGCTCAATTA 2140
 Db 138061 TTCTGTCTCCAGACTTAATCCCAATCTCTGCAAGGTGTCTGCTCGCTCAATTA 138002

QY 2141 GAACTGCGAAGAGTACTGAGGCTGCGACAGTACTAGAGTAAAGAAAGTAACTA 2200
Db 138001 AGAATTTGGCAAGGGGTGCAAAACATTTGCAG----- 137969
QY 2201 GCACCTCTAGAGAGCAAGAAAATGGGCAACCTTCATGATGCTGCTTCCCTTCT 2260
Db 137968 -----TATTTGGAAAACAACAAAAGTCAACCTTTATCATGATGCTCTTCTCTTTAT 137915
QY 2261 GTGTTCACATGCTCCGATATATTTAAGAGAGGTAGATA-----ACTACATTTT 2312
Db 137914 GTGTTCATATATTTCTGATATATTTATAGAGAAATAGTATGCACTTTTATCTCTCG 137855
QY 2313 TCTTTTACCTGGAAGGCTGAGAAAACCTTTGTTACCATATATAAATTCATCTTC 2372
Db 137854 ATATTACTGCTGGAAATCTGAGGCAACCTGTAATAGCTGCTGCATCCAGTTATATAA 137795
QY 2373 TAACTGCTATGTTATTTCTAAGATCAATGATGAGATATCTCTTTGTTATATAA 2432
Db 137794 TTTATTTATCTATGTTCAAGATTTTCTACTAGCTGACATTAACCTCTTGTGTTATAA 137735
QY 2433 ATGAAAAAACAACATCTCTGAGCATATTTATCTGCACTTTAGAGTAACTTA 2492
Db 137734 ATG-AAAAACATCTTCTGATTTATGTTAATCTGCACTTTAGATAGAA-----ATA 137681
QY 2493 ATACTAGTCAATGAACTGAAATCAATTTTCATATGATATAAGATAT-----TATTT 2547
Db 137680 ATAGCACTCAGTCAAAAGTTCTGATTAATTTTCATATTAATAAGACATTAATATGAT 137621
QY 2548 AAAAGTATTCATGAGCAATTTATATTA--AAGTATGATTTCTGTTATGTTAAAGAT 2606
Db 137620 AAAATATATTCATGAGCAATTTATATTAATAATACTTCCAAATTTTACTTAAGAT 137561
QY 2607 TTAATCAGGGAACAAGTTTCTCAATTTATAGCAAAATCTTTACTAGTACAGCTC 2666
Db 137560 TTAGTCA-GAAAAACAAGTTTCTCAATTTATAGATGAAATCTCACTAGTATCATCTC 137502
QY 2667 TTTTCATTTAAGTC-TTCCGATATATCTGATTTTCTAATTTATACAAAGCTAAATA 2725
Db 137501 TTAATCTTTAATTCAGGCTTCTCAATTTTATTTTCTTAATTTACTGGCACTAAATA 137442
QY 2726 ATTTAATTAACAATAATTAATTTATTT--TACTGAAAGCTTACTTAATGTTAAAT 2784
Db 137441 ATTTAATTAACAATAATTAATTTATTTCTACTCAAAATCTGCTTAATAACATTAATAAT 137382
QY 2785 CATTATATTTCTGAGGAAAAAGCATATCACTTTTATAGTATGATGATTAATAGAT 2844
Db 137381 CATTATATTTTGAAGAGTATATTTTC-ATATTTTAAATATGATATTAATTAATAAT 137323
QY 2845 TTATTCATTAATTAATTTTAAAGTTTCAATATTAAGATTAAGATCTAAGTG 2904
Db 137322 TGACTTATTTAATTAACAATTAAGTGTGAGAGATTAAGATTAAG-----TACAG 137267
QY 2905 TATATGTTATGTTAATAAGTTTATTTTATTTTGGCAGTCAATACAGCCCTTATTTAT 2964
Db 137266 TTTATATTAATGCAACATAGCTTTTGTGTTTCAAAATGTCACAAATTAACCTTTATTTAT 137207
QY 2965 CATAGATTTATCTTTTAAAGAGTATCAATGATGATGATGATGATGATGATGATGAT 3024
Db 137206 TGTGATTTATTTCTTTTATAGAGTATCAATGATGATGATGATGATGATGATGATGAT 137147
QY 3025 AATATCTGTTGGTGACTTTGACAGAGGCTTTTAACTCTGACAGCGAATTCATTTGT 3084
Db 137146 AATATCTGTTGGTGACTTTGACAGAGGCTTTTAACTCTGACAGCGAATTCATTTGT 137087
QY 3085 GGAGCAAGAGCAATCAAGATCCGAGAGCACTGTTCTCATCAAGTTGGAATTAATAA 3144
Db 137086 GGAGCAAGAGCAATCAAGATCCGAGAGCACTGTTCTCATCAAGTTGGAATTAATAA 137027
QY 3145 AGCACTTGAATACAGTATTAAGATTCAGTGTGTGGCAAGTTGTCTCTCAGAGCTGG 3204
Db 137026 AGCACTTGAATACAGTATTAAGATTCAGTGTGTGGCAAGTTGTCTCTCAGAGCTGA 136967

QY 3205 CAGCATTAACCTTTGGCTTGCGCTTACTCAAAAAGCAAAAAGTAAAGAGAACT 3264
Db 136966 CATCATTAATAATTTGTTGGCATTTCTCAAAAAGCAAAAAGTAAAGAGAAAC 136907
QY 3265 AAGAAAGAGTGTATGATTTTAAACCATGCAAAACATGCAATCTCTGT 3323
Db 136906 AAGAAAGAGTGTATGATTTTAAACCATGCAAAACATGCAATCTCTGT 136847
QY 3324 TTAATTTTACCTATTTATGCTGATTTGCTGCGCCAGTGGATGATGATGATGAT 3383
Db 136846 TTAATTTTACCTGTTATGCTGATTTGCTGCGCCAGTGGATGATGATGATGATGAT 136787
QY 3384 GCAGAA 3389
Db 136786 GCAAAA 136781
RESULT 12
AR211435
LOCUS AR211435 3435 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6399312.
ACCESSION AR211435
VERSION AR211435.1 GI:2514756
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3435)
AUTHORS Wu-Mong J.R. and Wang J.
TITLE Myostatin gene promoter and inhibition of activation thereof
JOURNAL Patent: US 6399312-A 1 04-JUN-2002;
FEATURES
Source 1..3435
/organism="Unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 43.0%; Score 1459.2; DB 6; Length 3435;
Best Local Similarity 72.7%; Pred. No. 1e-247;
Matches 2472; Conservative 0; Mismatches 743; Indels 187; Gaps 38;
QY 1 TGCCTGCCGAAGCATAGAAAGCTTGCATATCCCGCACTGAGAGGAAACACCTTC 60
Db 119 TTTATTTCTGGGCTCGAGAAAGTTGCATCACTGCACTGAGAGGAAATTAACCTTA 178
QY 61 TTGAATTTCTGAGTACTCTTATTCATTCATGCTGACCTCGGA-----GATTGG 114
Db 179 CTTAAATTTCTGAGTACTCTTATTCATTCATGCTGACCTCGGA-----GATTGG 238
QY 115 ATAGAGTGACTCTCATTTATTCAGAGGTTATGTTTACCCAAATCACTGCAACA-TGA 173
Db 239 ATATAGTCAATTTCTCCATATTCATAGTATGTTATGTTATTAAGTCAATGCAACATG 298
QY 174 ATATAGTATTCCTGAACCACTGCCCCCTAGAGGAACTCAAGGTATGATTTCCCGACCT 233
Db 299 GTTAGCAAAATTTGAACCTATTTGTTCCAGAGGAAACAGGGTATGAGGCTCTG-- 355
QY 234 CTGTGCAAGTTTGTATTAACCAATCAATTAATACTTTGTTTGTGCAATTTCTGTTT 293
Db 356 GTATTAATTTATATGATCAACCAATCAATTAATACTTTGTTTGTGCAATTTCTGTTT 415
QY 294 AAAATATCTTTATTTAATGATGATGCTAATTTCTTCAACATTTGTTTCAACCAAAAG 353
Db 416 CAATTA-CATTTATTAATTAATTAATTTGTTTCAATTTCACTGAACCTGACAGCA---G 470
QY 354 CCCATTAATCTGAGAGCCCTGAATGAAGTCTTACATAACACATGTTTTCATATGAGA 413
Db 471 CAGCACTTAATCTGAGAGCCCTGAATGAAGTCTTATCTTACACA---TGATTTTTCATGAG 527
QY 414 AAAATTTTCTTCAAGTCTGCGACAGCTTCTGCTTAATAATTTGGAACAAATATACA 473
Db 528 AACTTTTCTTCAAGTCTGCGACAGCTTCTGCTTAATAATTTGGAACAAATATACA 582

474 TAACTGAAAGTACTATATTAATGATTTTAACTGTACAGTTCAAGTCATTAGTA 533
583 AACTTATGACA-----TCTTAACTATTTAGGTGATGAGTTCCGGTGGCATTAAGCA 633
534 CACTTACCTTGTGTGGAACCTATGATCAACATTTCACTCTAGAACCTTTTGTATCTTCTG 593
634 CATTCAACTGTGTGTGATCAATCAACACATTTCACTTCAAGAAATTTTTCATCTTCCCA 693
594 AATTGAAACTCTGTGATCTATGACAGTAGCTTTCCCATCAACCGACACCTTCCCTC 653
694 GACTG-AAACTCTGTATCTATCAACAGTAATCTACCTCTCA-----GTGCTCAAC 743
654 CAGCCCCCGGACCAACCATCTTCTTTCATGACAGTCATCTGTGCTTAGAAC- 710
744 CAGTCCCTGGCAACCAACCATGCTACTTTCATTTGACAGCTTCTGTGATTAGAACATAG 803
711 -----CAGCCTTCAACATACCTTGTGGGGGCACTGT---AAGCAACGAGATC 756
804 GAACACACGACGACGTCAGACATATGCTGGGGGCCATTTTAAAGCAAAATTCAT 863
757 ACTCCCTACCGCACCAATGACACAAAATATATAAAGCATGGTGGCATTCGATTCG 816
864 AAGAGAGACATTAATAAAGAAAGACCAAAATATGTAAACATGGCACTAAATTAACGG 923
817 AAAAAAGGTCTGCT--AAGTAGAGGGCTGAAAC---AAGCAGAGAAATTAAGTAGT 871
924 AAAAAAGGAATTTGTTATATGATAGAGCTGAACAAAAGCGGAACCTGCTGTGT 983
872 TGACCTCAGCTGGATCTCTGTGTGTG--AAGCTCAAAATTTTCATTTGTTCTGTGAT 929
984 TGACCTCAGCTGGGAACCTGTGCTGATAGATGACCTCAAAATTTTGTGCTGTGTGAT 1043
930 ACGCAAAATGCTTATAAAGCACTGTAAAGATGATTAAGTAAGTAAATCTGAC 989
1044 GCCAGCAATGACTAAAA---TGTAGGAATTTGAGTTTCAATAGTAATTTTAC 1099
990 CAAGCATTAATGTGCAAGCAACGAGATCCATGATTAAGACAGCATGACAT--GTGAAA 1047
1100 AAGTAGGAATTTGTGATACAGAAATCATGATTAATTAAGAAATTTGTCTGTTAGGGAAA 1159
1048 TGATATCTTTGTTCTTTATTTCAAGGACGTAAAGAAAGCG-----T 1093
1160 TGACAACTTTGTTCACTTTTCCAGGTGATTAAGAAATTAAGAGCTGAAACATGTCTGT 1219
1094 CACAGGGCTGCTTACACATTTTACTAGAGCTAGGCTATGTGAGTGG----- 1144
1220 AACTGGCTGCTTGAATAATTTTACTTGTCTCAAGGCTAGAAAGTAGGCTTTGTGAC 1279
1145 ---TAGTGGCAATTACAACACTGAAGAGTTCTAGTTGATGTGAGATGATTTTACCA 1201
1280 TGATTAATTGGAGCTATGACCTGAAGAGTTCTAGTTGATGTGAGATTAATTTTACCA 1339
1282 TAATTTCAACCCCTCTGTGATGAACAGAGACTAAGTACTCAAGTACAGATTATGATCA 1261
1340 TAATCTCAACCCCTCTGTGATGAACAAAGAGCAAGCACTCAAAAGCCAGATTATCAATTA 1399
1262 CTACTATATGACAGGACCTGACTCAACAATTTTACATGATTAATTTTGAATTAATGAGCC 1321
1400 CTACTATATGACAGGCTGACTATTAAGCAATTTTACATGATTAATTTTATATATCCCC 1459
1322 CAACACTATAGAGAGCTGAAGGTTAG-----AAGTATCTCATTTATTTATACCA 1376
1460 CAAAACCTATAGAGAGCTAAAGTTTAGGAAGTTAAGTATCTATCATTTATTAACATA 1519
1377 GT-----GGCAACTGAGATCTGAACCTGCTTATCCAACTCCAGAGCTGAGATC 1428
1520 GTTAGAAGTGGCAAGTTGAGATTTGAACCTGAGTCTATCTGATCCAGAGCTGAGTTC 1579
1429 CCAATTGTACACATTTCAATCAAGTTAAAGGAAAAGGATTTGATTTGCTGAGAG 1488
1580 TCAATTTCAACTGCTATACATTTTAAGCATTTATAAAGAAAGTTGACTTACCTGGAAC 1639
1489 TGTATAGGGCATATGTTACA---ATTATTAATTACAAAGATTTATATGTTGAAAATA 1545

1640 TGTATAGATGATGTGTTACATGATCATTAACATTTGAAAATTTACATGAAAATG 1699
1546 AATTATCAACAAATTAACCTTATTAAGCCGTATCTAATCTGCTCGGCAACAAAGCT 1605
1700 AATTATCAACAAATTAACCTTATTAAGCCGTATCTAATCTGCTCGGCAACAAAGCT 1759
1606 ATCTGAATCCCTGAGGGCATCTGGTTGTGTCTGTTTCTTAACTTAAATGATAGG 1665
1760 ATCTGAATCC--TGAAGGCATCTGTTGTGTGTCTGTTTCTTAACTTAAATGATAG 1818
1666 CAATCTAATGATTAATGTAAGGCA--TTTTTCTCAAGAGATGATATCTTAAAGA 1724
1819 CAATCTAATGATTAATGTAAGGCA--TTTTTCTCAAGAGATGATATCTTAAAGA 1878
1725 AATTGTAAGAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1784
1879 AATTGTAAGAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1938
1785 GTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1841
1939 GTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1998
1842 ACCACTGACTCTCTTGG--GAAAAAAGTAATGATTAATGATTAATGATTAATGATTA 1900
1999 GCACTGATTTCTCTTGGAAAAAAGTAATGATTAATGATTAATGATTAATGATTAATG 2058
1901 TCCCAAAAGTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1960
2059 TCCCAAAAGTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2118
1961 CTGGCTATTAATTTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2020
2119 CTGGCTATTAATTTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2178
2021 CTAAAGTTGAGATTTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2080
2179 CTAAAGTTGAGATTTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2238
2081 TTCTCTGCTCCCAACCTTACCCCAATCTCTGCAAGTGTCTGCTCCGCTGCAATGA 2140
2239 TTCTCTGCTCCCAACCTTACCCCAATCTCTGCAAGTGTCTGCTCCGCTGCAATGA 2298
2141 GAACTGGCAAGGAAGTACTAGAGAGTCCGACAGTACTAGAGGAATGAATGACCTA 2200
2239 AGAATTTGGCAAGGGGTGCAAC-----TA 2325
2201 GCACACTGAGAGAGCAAGAAAATGGGACCTTCAATGATGTGCTTCCCTTCT 2260
2326 TCGAGTATTTGGAGAAACAAAGAGTCAACCTTTATCATGATGCTCTTCTCTTAT 2385
2261 GTTCTCAATGCTCCCATTAATTAATTTACAGAGGTAATA-----ACTACATTTT 2312
2386 GTGCTCAATTAATTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2445
2313 TCTTTTACCACTGGAAGCTGAGAAAATCTTGTGTAACCATCAATTAATTAATTAATTA 2372
2446 ATATTTACTGTGGAATCTGAGCAACCTTAATTAATCTGCAATGCACTTAATTA 2505
2373 TAAGCATCTTATGTTATTTTAAAG-----TCAATTAAGTGAACATTAATCTTGTGTA 2426
2506 TTCAATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2565
2427 TAAACATGAAGAAACACATCTCTGAGCAATTAATTTACCACTTAAAGATAGAGT 2486
2566 TAAACATGA--AAAAACATCTCTGAGTATGTTAATCTGCACTTAAAGATAGAG-- 2622
2487 AACTTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 2542
2623 ---ATAATGACATGACCAAGTCAAGTATTAATTTTCAATTAATTAATTAATGACATG 2679
2543 -TATTAAGGAATTTCCATGAGCAATTAATTAATTA--AAGTAGAATTTTCAATTAATG 2600

Qy 990 CAAGACATAAATGTCAGACGCGGATCCATGAATTAAGACACTGACCAT--GTGGAAA 1047
Db 1100 AAGTAGCAAAATTTGTGAATACAGAAATCATGAATTAATGAATTTGTCTGTAGGCGGAAA 1159
Qy 1048 TGAATATCTTTGTTTCTTTATTCAGAGCACTAAGAGAAAGCCG-----T 1093
Db 1160 TGACAACTTTGTTCACTTTTTCAGGTGATGAATAAGAGCTGGAACATGCTTGT 1219
Qy 1094 CACAGGGCTGCTTAACACATTTTACTAGAGCTAGGCTATGTCAGTCGG----- 1144
Db 1220 AACCTGGCTGCTTAAGAAATTTTACTTCTCAGAGGCTTAAGAAAGTAGGCTTTGTAC 1279
Qy 1145 ---TAGCTGGCAATTAACAACCTGAAGCACTTACTGATGTGAGATGAATTAACCA 1201
Db 1280 TGATTAATGGCAGCTATGACCTGAAGCAGTTCTAGTTCATGTGGAGCAATTAATTAACA 1339
Qy 1202 TAAATCTCAACCCCTCTGCTGAGAAACAGAGACTAAGTACTCAAGTACAGATTATGATCA 1261
Db 1340 TAAATCTCAACCCCTCTGCTGATMAAACAAAGAGCAACCTCAAAATGCCAGTATCAATTA 1399
Qy 1262 CTTAATCTATGACAGGCACTGATCAACATTTTACATGATTAATGAATTAACATGCGCC 1321
Db 1400 CTTAATCTATGACAGGCTGATCAATTTACAGCAATTTACATGATTAATTAATTAATCCCC 1459
Qy 1322 CAAGACTCTATGAGAGAGCTGAAGGTTAGAG---AAGTATCTCATTCATTAATTAACA 1376
Db 1460 CAAGAACCTTATGAGAGAGCTAAAGTTTAGGGAAGTTAGTATCTCATCTATTAACATA 1519
Qy 1377 GT-----GGCAAACTGAGATCTGAATCTGAGTCTATCCAATCTCCAGACCTGAGATC 1428
Db 1520 GTTAGAAGTGGCAAGGTTGAGATCTGAGGCTATCTGACTCCAGAGCTGAGATTC 1579
Qy 1429 CCAATCTGATACACATTTCTAATCAAGTTAAAGGAAAAAGAGATTTGATTTGCTCAGAG 1488
Db 1580 TCAATTTCACTGCTATACAAATTTCTAGCAATTTAAAGAAAGTTGATTTGACTTGTGAC 1639
Qy 1489 TGTATAGGGGCAATGTACA---ATTATTAATTACAAAGATTTATATGTTGAAAAATA 1545
Db 1640 TGTATAGATGATGATGTTCATATGATCATTAACATTTGAAAGATTTACACATTTGAAAAATG 1699
Qy 1546 AATTTATCAAAACAATTAATCTTTATAGCTGATCTAATCTGCTCCGCAACAAAGCT 1605
Db 1700 AATTTATCAAAACAATTAATCTTTATAGAGCCAGATCTAATATTTGCCATTAACAAAGGT 1759
Qy 1606 ATCTGAATTCCTTACAGGCACTGCTGTTGTCTGCTGTTTCTCCTAATCTTTAATGATGAG 1665
Db 1760 ATCTGAATTC--TCAGGCACTCTGTTGTCTGCTGTTTCTCCTAATCTTTAATGATGAG 1818
Qy 1666 CAAATCTAATGATGATTAAGGCA--TTTTTCTCAAGAGATGATGATCTCTTAAGA 1724
Db 1819 CAAATCTAATGATGATTAAGGCAATTTTTTCTCAAGAGATGATGATCTCTTAAGA 1878
Qy 1725 AATTTGATGAAAAATGATTAACCTTTTCAAGGCTACTGAGTTGCAATTTAGTGACTGAGCA 1784
Db 1879 AATTTGATGAAAAATGATTAACCTTTTCAAGGCTACTGAGTTGCAATTTAGTGACTGAGCA 1938
Qy 1785 GTAAATTTGATGAAAAATGAGGAAA---GTAGTGACTTAAAGAAATTAATTTGATGA 1841
Db 1939 GTAAATTTGATGAAAAATGAGGAAAATGATGATGATTAATTAATTTGATGA 1998
Qy 1842 ACCATGCACTCTCTTGG--GGAAAAAGTAAATGATTAATCTCTTAAGAGTCTTACT 1900
Db 1999 GCCATGATTTCTCTTGGAAAAAAAGTAAATGATGATTAATCTCTTAAGAGTCTTACT 2058
Qy 1901 TCCCAAAAGATGAGAAAAATTAATCTCTGAGGCTGGAACAGCTTCTGTTCTTG 1960
Db 2059 TCCCAAAAGAGTGAAGAAAAAGAAATCTCTTGGCTAGAAATATCTTCTGTTCTTG 2118
Qy 1961 CTGGCTAATTTTGTAGTTTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2020
Db 2119 CTGGCTAATTTTGTAGTTTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 2178
Qy 2021 CTAAAGTTGAGATTTGATTCCTTACAGAGGCCACTTCAATTTAGAGAAACAAAGCCCCA 2080

Db 2179 CTAAGTTGAAAGCTTTGATCCCTTACAGAGGCCACTTAATTAATTAAGAGAAACAAAGCTCTTA 2238
Qy 2081 TTCTGTGCTCCAGACCTTATCCCAAAATCCCTGACAGTGTCTGCGCTCCGGTCAAAATGA 2140
Db 2239 TTCTGTGCTCCAGACCTTATCCCAAAATCCCTGACAGTGTCTGCGCTCCGGTCAAAATG 2298
Qy 2141 GAACTGGCAAGAGATCTAGAGAGGTGACAGTACTAGAAAGTAAAGAAATGAGACTA 2200
Db 2299 AGAAATGGCAAGAGGAGTGAACA-----TA 2325
Qy 2201 GCACATCTGAGAGAGAGAAAAATGGGACCCCTTCATGATGATGTTCTTCCCTTCT 2260
Db 2326 TCGCAGTATGGGAGAAACAAAGAGTCAACCCCTTTATCAAGATGCTCTTCTCTTTAT 2385
Qy 2261 GTGTCAAAATGCTCCGATATATTTTACAGAGGATGA-----ACTAATTTT 2312
Db 2386 GTGCTCAATTAATTTCTGATTAATTTATAGAAATGATGATGACATTTTACTCTCTGG 2445
Qy 2313 TCTTTTACACTGGAAGGCTGAGAAAACTTTGTTACCATCATMAAATTCATATCTTC 2372
Db 2446 ATATTTACTGCTGAAATCTGAGGCAATCTGATTAATCTGCTGACATGCGAGTTATAAA 2505
Qy 2373 TAACTATCTATGTTATTTCTAAGA-----TCAAAATGCTGAATATCTCTTTGTA 2426
Db 2506 TTCAATTTCTAATGCTATGTTTCAAGATTTTCTTACTAGCTGAGCATTAACCTCTTGATA 2565
Qy 2427 TAAACATGAAAAAACAATCTGAGCAATATTAATCTGCACTTTAGATGAGAGT 2486
Db 2566 TAAACATG--AAAAACATCTCTGAGTTATGTTAATCTGATCTTTTGAATGGA-- 2622
Qy 2487 AACTTAATCTAGTCAATTTGAACATGAATCAATTTTCATATGATTAATGAATAT-- 2542
Db 2633 ---ATAATGACATCGACAAAGTGTGATTAATTTTCAATTAATTAAGACATGAAA 2679
Qy 2543 -TATTTAAAGTAATTCATGAGCAATTAATTAATTA--AAGTGAATTTTCAATATGCTT 2600
Db 2680 CTATGTAATAATTAATTCATGACATGATTAATTAATTAATTAATTAATTAATTAATTA 2739
Qy 2601 AAGAAATTAATCAAGGAAACAAAGTTTCTCAATTAATGACAAAAATCTTTACTAGTATC 2660
Db 2740 AAGAAATTAATCA--GAAAAAGATTTCTCAATTAATGATGATGATTAATTTCTCACTGATATC 2798
Qy 2661 ACAGTCTTTTCAATTAAGTCT--TTCGTAATTAATCTGATTTTCTAATTAATTAACAAGCTA 2719
Db 2799 ATAAATCTTAATTTAATTAATTAATTAATTAATTTTCTAATTAATTAATTAATTAATTA 2858
Qy 2720 AAAATTAATTAATTAATTAATTAATTAATTTT--TACTTAATGCTTACTTAATTAATGA 2778
Db 2859 AAAATTAATTAATTAATTAATTAATTAATTTTCTAATTAATTAATTAATTAATTAATTA 2918
Qy 2779 TAAATATCTTTATTTTCTGAGGAGAAAGCATATCACTTTTAAAGTATGAAGTAAAT 2838
Db 2919 TAAATATCTTTATTTTGAAGAGTAAATTTTCT--AAATTTTAAATTAATTAATTAATTA 2977
Qy 2839 TAAATATCTTTATTTTGAAGAGTAAATTTTAAAGTATCAATTAATTAATTAATTAATTA 2998
Db 2978 TAAATATCTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3037
Qy 2899 TAAATATCTTTATTTTAAATTAATTAATTTTAAATTTTTCATGATGCAATTAACGCTTT 2958
Db 3038 ---AGTTATTAATTAATTAATTAATTAATTTTTCATTAATTAATTAATTAATTAATTA 3093
Qy 2959 ATTAATCTAATTAATTTATTTTAAAGAGTAAATTAATTAATTAATTAATTAATTAATTA 3018
Db 3094 ATTAATCTAATTAATTTATTTTAAAGAGTAAATTAATTAATTAATTAATTAATTAATTA 3153
Qy 3019 TAACAAAATCTGTTTGTGATGATGAGAGAGGTTTAACTCTGACAGAGAT 3078
Db 3154 TAACAAAATCTGTTTGTGATGATGAGAGAGGTTTAACTCTGACAGAGAT 3213
Qy 3079 CATTTGGAGCAAGGCCAATCAAGATCCGAGAGCACTTGTCTCATCAAGTTGAAT 3138

Db 3214 CATTGTGAGCAGACGCCAATCATATGATCTGACGACACTGTCTCATCTAAGTTGGAA 3273
Qy 3139 ATAAAAAGCCACTTGGAAATAGATATTAAGATTCTAGTGTGGCAAGTTGTCTCTAG 3198
Db 3274 ATAAAAAGCCACTTGGAAATAGATATTAAGATTCTAGTGTGGCAAGTTGTCTCTAG 3333
Qy 3199 ACTGGGCGAGGCACTTACCTTTGGCTGGCGCTTACTCAAAAGCAAAAAGTAAAGGA 3258
Db 3334 ACTGTACATGATTAATTAATTTTGTGGCATTAAGCAAAAGCAAAAAGTAAAGGA 3393
Qy 3259 AGAGTAGAACAAGGAAAGATTGTATGATTTTAAACC 3300
Db 3394 AGAACAAGAACAAAGAAAGATTATATGATTTTAAATC 3435

RESULT 14
BTA38578
LOCUS BTA38578 1269 bp DNA linear MAM 08-APR-2002
DEFINITION Bos taurus partial gdf-8 for myostatin and promoter region.
ACCESSION AJ438578 GI:20135841
VERSION AJ438578.1 GI:20135841
KEYWORDS gdf-8; myostatin.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE 1
AUTHORS Crisa, A., Marchitelli, C., Savarese, M. and Valentini, A.
TITLE Sequence analysis of the myostatin promoter region in cattle
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1269)
AUTHORS Crisa, A.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2002) Crisa A., Produzioni Animali, Università degli Studi della Tuscia, Via S. Camillo de Lellis s.n.c. 01100 Viterbo, ITALY
FEATURES
SOURCE Location/Qualifiers
1..1269
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
gene complement(237..242)
/gene="gdf-8"
enhancer complement(237..242)
/gene="gdf-8"
gene 1008..1269
/gene="gdf-8"
CAAT_signal 1008..1011
TATA_signal 1051..1057
/gene="gdf-8"
TATA_signal 1075..1081
/gene="gdf-8"
CDS 1214..>1269
/gene="gdf-8"
/codon_start=1
/product="myostatin"
/protein_id="CAD27441.1"
/db_xref="GI:20135842"
/db_xref="UniProt/TrEMBL:Q8SPM3"
/translation="MQKQISVYLYLFLMLIVAG"

ORIGIN
Query Match 36.9%; Score 1251.2; DB 4; Length 1269;
Best Local Similarity 99.5%; Pred. No. 6.2e-211;
Matches 1264; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 2087 GCTCCAGACCTTACCCCAATCCCTGCAAGTGTCTGCTCCGCTCAATGAGAAGT 2146
Db 1 GCTCCAGACCTTACCCCAATCCCTGCAAGTGTCTGCTCCGCTCAATGAGAAGT 59
Qy 2147 GGCAGAGAGAGTACTAGAGGTCCGACAGTACTAGAGAGTAAAGTGAAGTGAAGT 2206

Db 60 GGCAGAGAGTACTAGAGGTCCGACAGTACTAGAGAGTAAAGTGAAGTGAAGTGAAGT 119
Qy 2207 TACTGAG 2266
Db 120 TACTGAG 179
Qy 2267 ACAATGCTCCGATTAATTTTACAGAGGATAGATTAATTTTCTTTTACCACTG 2326
Db 180 ACAATGCTCCGATTAATTTTACAGAGGATAGATTAATTTTCTTTTACCACTG 239
Qy 2327 AAGCTGAG 2386
Db 240 AAGCTGAG 299
Qy 2387 TTAATTAAGTCAATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2446
Db 300 TTAATTAAGTCAATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 359
Qy 2447 CCTCTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2506
Db 360 CCTCTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 419
Qy 2507 AAATGAAATCAATTTTCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2566
Db 420 AAATGAAATCAATTTTCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 479
Qy 2567 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2626
Db 480 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
Qy 2627 CTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2686
Db 540 CTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 599
Qy 2687 AATTAATCTGATTTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2746
Db 600 AATTAATCTGATTTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 659
Qy 2747 TATTTTACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2806
Db 660 TATTTTACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 719
Qy 2807 GCATATCACTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2866
Db 720 GCATATCACTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 779
Qy 2867 TTAAGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2926
Db 780 TTAAGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839
Qy 2927 TTTTAAATTTTTCGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2986
Db 840 TTTTAAATTTTTCGAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 899
Qy 2987 AGTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3046
Db 900 AGTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 959
Qy 3047 CAGACAGGTTTAACTTCTGACAGGAGATTAATTAATTAATTAATTAATTAATTAATTA 3106
Db 960 CAGACAGGTTTAACTTCTGACAGGAGATTAATTAATTAATTAATTAATTAATTAATTA 1019
Qy 3107 CCGAGAGACCTGTCTCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3166
Db 1020 CCGAGAGACCTGTCTCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1079
Qy 3167 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3226
Db 1080 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1139
Qy 3227 CGTTACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3286

```

Db      1140 CGTACTCAAAAGCAAAAGAAAGTAAAGAGAGAGTAAAGACAGGAAAGATTGTA 1199
Qy      3287 TTGATTTTAAACATGCAAAACTGCAAAATCTGTTTATTTATTAATTTATGCGA 3346
Db      1200 TTGATTTTAAACATGCAAAAGTCAAACTGCAAAATCTGTTTATTTATTAATTTATGCTGA 1259
Qy      3347 TTGTTGCTGG 3356
Db      1260 TTGTTGCTGG 1269

RESULT 15
AC140969/c
LOCUS
DEFINITION Canis familiaris clone RP81-197A18, WORKING DRAFT SEQUENCE, 11
ordered pieces.
AC140969 192219 bp DNA linear HTG 16-MAY-2003
AC140969.2 GI:30794531
HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE
Canis familiaris (dog)
ORGANISM
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 192219)
Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooke,S., Carliaga,K.,
Chu,G., Coleman,B., Coleman,H., Enogle,J., Granite,S., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurtle,B., Idol,J.R., Karlins,P., Kwong,P., Latic,P., Lee-Lin,S.-Q.,
Legaspi,R., Maduro,Q.T., Maduro,V.B., Margulies,E.H., Mastello,C.,
Maskeji,B., McDownell,J., Paguirigan,C., Pearson,R., Portnoy,M.B.,
Prasad,A., Reddi-Bughe,N., Schandler,K., Schueler,M.G., Shah,K.,
Sison,C., Stancirup,S., Thomas,J.W., Thomas,P.J., Tidipuri,V.,
Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 192219)
Green,E.D.
Direct Submission
3 (bases 1 to 192219)
Submitted (06-MAR-2003) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
Green,E.D.
Direct Submission
Submitted (16-MAY-2003) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
On May 16, 2003 this sequence version replaced gi:28867001.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center Project name: eaj
Center clone name: 197A18

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
AUTHORS
JOURNAL
COMMENT

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 190062 bases at least Q40

Consensus quality: 190807 bases at least Q30
 Consensus quality: 191127 bases at least Q20
 Insert size: 172000; agarose-fp
 Insert size: 191219; sum-of-contigs
 Quality coverage: 10.69x in Q20 bases; agarose-fp
 Quality coverage: 9.61x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1927: contig of 1927 bp in length
1928 2027: gap of unknown length
2028 88251: contig of 86224 bp in length
88252 88351: gap of unknown length
88352 107063: contig of 18712 bp in length
107064 107163: gap of unknown length
107164 109235: contig of 2072 bp in length
109236 109335: gap of unknown length
109336 119132: contig of 9797 bp in length
119133 119232: gap of unknown length
119233 125483: contig of 6251 bp in length
125484 125583: gap of unknown length
125584 140954: contig of 15371 bp in length
140955 141054: gap of unknown length
141055 143181: contig of 2127 bp in length
143182 143281: gap of unknown length
143282 158769: contig of 15486 bp in length
158770 158869: gap of unknown length
158870 163849: contig of 4980 bp in length
163850 163949: gap of unknown length
163950 192219: contig of 28270 bp in length.

```

FEATURES

source

```

1..1927
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
/clone="RP81-197A18"
/clone_1bp="RP81"
1..1927
/feature="assembly_fragment"
clone_end:17
vector_side:left"
2028..88251
/feature="assembly_fragment"
88352..107063
/feature="assembly_fragment"
107164..109235
/feature="assembly_fragment"
109336..119132
/feature="assembly_fragment"
119233..192219
/feature="assembly_fragment"
AC140967 clone RP81-292C20 (center project name eak)"
119233..125483
/feature="assembly_fragment"
125584..140954
/feature="assembly_fragment"
141055..143181
/feature="assembly_fragment"
143282..158769
/feature="assembly_fragment"
158870..163849
/feature="assembly_fragment"
163950..192219
/feature="assembly_fragment"
clone_end:SP6
vector_side:right"

```



```

QY 2280 -----ATPA 2283
Db 90638 AAAATAAGGCTTTCTGATTTCTCAAAATAGAAATGATCTTAAAAATAATATATPA 90579
QY 2284 TTTACAGAGGTGATATACATTTTCTTTTACACCTGGAAGCTGAGGAAACTT 2343
Db 90578 CTTATGAGAGAGAT-ACATAGCTTTTCTTTTACCTGTAGAAATCTGAGGAAAGTG 90520
QY 2344 T-----GTACCAATCATAAATTTCACTATCTTCTAAGTCAATCTATGTATT 2391
Db 90519 TAGCAAGTGCAGATTTGTACGTATACAGTATCTTCTTAT-TCATAGTTGAGGTTT 90461
QY 2392 CTAAAGTCAATAGCTGACAAATATCTCTTTGTAATAACATGAAAAACACATCTCT 2451
Db 90460 TTTTCTCTAGATGACATTTAGCTCTTGATATACATGAAAGAAACAGAACTTCT 90401
QY 2452 GAGCAATTTATCTGCACTTTAGATAGGAAGTAACTTATCTAGTCAATTTGAACT 2511
Db 90400 GAGTAAATTTAACCTGCACTTCAAAATAGGAATTAAGTGAATGCAATCAGCCAAAAG 90341
QY 2512 GAAATCAATTTTCAATGATATAA-----GATATATTAAAAAGTAAATTCATGAGCA 2566
Db 90340 TGAGCAGAGTTTCAATTTATATAAACAATGAAATTTTCAAAAGAAATTCATGTGTA 90281
QY 2567 ATTTAATATT--AAAGTAGATTTTCAATATGTATAGATTTATTCAGGAAACAG 2623
Db 90280 ATGAAACATTTGTAAGAGTTGATTTTAGATGTGTAAGATTTATTCAGGAAACAG 90221
QY 2624 TTTCTCAATTTATGCGAAATTTTCTAGTATCA-----GTCTTTTCAT 2673
Db 90220 TTTCTCAGATTAATAGCGAAATCTTTTACAGTACAAATGTCACTTTTATTAT 90161
QY 2674 TTAAGTCTCTGTAAT--AAATCTGATTTTCTAATTTATCAAGCTAAATTAATTTAA 2731
Db 90160 TCAGGTTCTCTTATTTCAAAATTCACAGTCCCTAATTTATACACATTAACATTA 90101
QY 2732 TATAACAATTAATTTTCTTCAATGCTTACTTAATATGATATAAATCATTTTA 2791
Db 90100 TGCAATTAACAATTTCTTTTACTTAATGCTTGCCAAATCATAGACAATTAATTTTA 90041
QY 2792 TTTTCTGAGGAAAGCATTCATCTTTTAAGTATGAGTAAATTAAGTTATTCA 2851
Db 90040 ACTTTAGAG--AAGTAACTTAACATTCAGTTTAAAGTGTGAGCAACAAATTTACTTA 89984
QY 2852 CTTAAATTTATATTTTAAAGTTTCAATTAAGATGAATGAATGAATGAATGAATGAATGA 2908
Db 89983 TTTTAATCACACTT---GGTTTCAATATGAAATGAATGAATGAATGAATGAATGAATGA 89927
QY 2909 --TGTATGTATATAAAGTTTAAATTTTTCGATGTACATACAGCTTTATATTCA 2966
Db 89926 TCTGTTATGTATACCAAGCTTTAAATTTTCAATGTACATATATCTTTATATTG 89867
QY 2967 TAGATTTATCTTTTAAAGATGATCAATGATCACTGACCTTGACTGTAAACAAA 3026
Db 89866 CGGATTTATTTCTTTTAAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 89807
QY 3027 TACTGTTTGTGATCTTGACAGACAGAGGTTTAACTCTGACAGAGATTTCACTGTTG 3086
Db 89806 TACTGTTGAGTCTTGACAGAGAGGAGGTTTAACTCTGACAGAGATTTCACTGTTG 89747
QY 3087 AGCAAGAGCAATCACAGATCCGACAGCACTTGTCTCATCAAGTTGAATATAAAG 3146
Db 89746 AGGAGAGCCATCATAGATCTGACAGACTTGTCTCATCAAGTTGAATATAAAG 89687
QY 3147 CCACTTGAATACAGTATAAAGATTCATCTGTGTGCAAGTTGTCTCTCAGACTGAGCA 3206
Db 89686 CCACTTGAATACAGTATAAAGATTCATCTGTGTGCAAGTTGTCTCTCAGACTGTA 89627
QY 3207 GGCATTAAGTTGGCTTGAGCTTCACTCAAAAGCAAAAGAAAGTAAAGAAAGAAAGTAA 3266
Db 89626 GGCCTTGAATTTGCTTGAGTCTCAAAAGCAAAAGAAAGAAAGAAAGTAA 89567
QY 3267 GAAACAAGGAAAGATTTGATTTGATTTTAAACCATGCAAAACTGCAAACTCTGTGTTA 3326

```

```

Db 89566 GAAAGAGGAAAGAGTGTCTGACTTTAAATCATGACAGACTGCAAACTGTGTTTA 89507
QY 3327 TATTTACCTATTATGCTGATGTTGCTGGCCAGTGTGATCTGAATGAGAACAGCAGCA 3386
Db 89506 TATTTACCTATTGCTGATGTTGCTGGCCAGTGTGATCTGAATGAGAACAGTGA 89447
QY 3387 GAA 3389
Db 89446 AAA 89444

```

Search completed: March 31, 2005, 12:46:14
 Job time : 9658 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 07:28:21 ; Search time 1147 Seconds
(without alignments)
17496.010 Million cell updates/sec

Title: US-09-743-312d-1_COPY_7103_10492
Perfect score: 3390
Sequence: 1 tgcgtccgcaagccatag.....atcgagacagcgagcaag 3390

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2001bs:*
6: geneseq2002s:*
7: geneseq2002bs:*
8: geneseq2003s:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3390	100.0	10492	3	AAZ58843	AAZ58843 Nucleotide
2	1462.2	43.1	3438	4	AAAC91773	AAAC91773 Human myo
3	525	15.5	2482	4	AAAF58661	AAAF58661 Murine my
4	508.6	15.0	5790	2	AAAX24464	AAAX24464 Bovine my
5	494.8	14.6	1537	4	AAAF58662	AAAF58662 Porcine m
6	422.4	12.5	649	3	AAAZ50110	AAAZ50110 Human gro
7	332.8	9.8	1537	4	AAAF58662	AAAF58662 Porcine m
8	259.4	7.7	396	3	AAAZ50113	AAAZ50113 Mouse gro
9	190	5.6	2823	12	ADH11572	ADH11572 Human bon
10	174.4	5.1	200	4	AAAF58663	AAAF58663 Murine my
11	141.8	4.2	156	3	AAAZ50115	AAAZ50115 pig growt
12	136.2	4.0	336	2	AAAV89102	AAAV89102 EST clone
13	127.6	3.8	1195	2	AAAX24415	AAAX24415 Bovine my
14	126.2	3.7	653	7	ADDS31200	ADDS31200 Human gen
15	126.2	3.7	2742	2	AAZ064449	AAZ064449 Genomic D
16	126.2	3.7	2743	2	AAV45813	AAV45813 Human gro
17	126.2	3.7	2743	2	AAZ09365	AAZ09365 Human GDF
18	126.2	3.7	2743	4	AAAF63550	AAAF63550 Human GDF
19	126.2	3.7	2743	6	ABK15393	ABK15393 cDNA enco
20	126.2	3.7	2743	6	AAAD29742	AAAD29742 Human pro

21	126.2	3.7	2743	10	ABQ76696	ABQ76696 Human pro
22	126.2	3.7	2743	12	ADM16470	ADM16470 Human pro
23	126.2	3.7	2743	12	ADM16493	ADM16493 Human pro
24	126	3.7	1239	2	AAAX24416	AAAX24416 Bovine my
25	124.6	3.7	2743	2	AAQ76372	AAQ76372 Human gro
26	118.4	3.5	799	3	AAZ50114	AAZ50114 Chicken g
27	115.2	3.4	564	2	AAV89872	AAV89872 EST clone
28	111.8	3.3	749	4	AAH04272	AAH04272 Human CDN
29	111.8	3.3	2277	4	AAH15506	AAH15506 Human CDN
30	104.6	3.1	2100	3	AAAG0289	AAAG0289 Mouse wil
31	104.6	3.1	2171	6	AAAD29743	AAAD29743 Murine pr
32	104.6	3.1	2675	2	AAZ06448	AAZ06448 Genomic D
33	104.6	3.1	2676	2	AAQ76371	AAQ76371 Mouse gro
34	104.6	3.1	2676	2	AAV42113	AAV42113 Murine gr
35	104.6	3.1	2676	2	AAZ09364	AAZ09364 Murine GD
36	104.6	3.1	2676	4	AAAF63549	AAAF63549 Murine GD
37	104.6	3.1	2676	6	ABK15394	ABK15394 cDNA enco
38	104.6	3.1	2676	8	ABK15456	ABK15456 cDNA enco
39	104.6	3.1	2676	10	ABQ76697	ABQ76697 Murine pr
40	104.4	3.1	330973	11	ACN44846	ACN44846 Human gen
41	103	3.0	2676	2	AAAX24417	AAAX24417 Mouse myo
42	102.8	3.0	355	4	AAI93466	AAI93466 Human pol
43	100.4	3.0	25000	12	ADJ34524	ADJ34524 Human GUC
44	96.2	2.8	348101	12	ADQ97146	ADQ97146 Human can
45	95.2	2.8	22475	12	ADQ97998	ADQ97998 Human can

ALIGNMENTS

RESULT 1	AAZ58843	standard; DNA; 10492 BP.
ID	AAZ58843	
XX	AAZ58843;	
XX	DT	25-APR-2000 (first entry)
XX	DE	Nucleotide sequence of bovine myostatin promoter-enhancer region.
KM	Myostatin; promoter; muscle cell disorder; genetic marker; bovine;	
KW	myogenic regulatory factor; oncogene; muscle growth;	
XX	muscular dystrophy gene; ss.	
XX	OS	Bos sp.
XX	Key	Location/Qualifiers
FH	enhancer	9094. .9099
FT	enhancer	/*tag= a
FT	enhancer	/note= "E-box 4"
FT	enhancer	9236. .9241
FT	enhancer	/*tag= b
FT	enhancer	/note= "E-box 3"
FT	enhancer	9860. .9865
FT	enhancer	/*tag= c
FT	enhancer	/note= "E-box 2"
FT	enhancer	9879. .9885
FT	enhancer	/*tag= d
FT	enhancer	/note= "TATA-box 3"
FT	enhancer	10095. .10100
FT	enhancer	/*tag= e
FT	enhancer	/note= "E-box 1"
FT	enhancer	10197. .10201
FT	enhancer	/*tag= f
FT	enhancer	/note= "CAAT-box"
FT	enhancer	10240. .10246
FT	enhancer	/*tag= g
FT	enhancer	/note= "TATA-box 2"
FT	enhancer	10264. .10270
FT	enhancer	/*tag= h
FT	enhancer	/note= "TATA-box 1"
FT	enhancer	10403. .10492
FT	enhancer	/*tag= i

QY 1621 GGGCATCTGGTTTGTGCTGGTTTCTTATCTTTATGATGGCAATCTAATGAT 1680
 DB 8723 GGGCATCTGGTTGTGCTGGTTTCTTATCTTTATGATGGCAATCTAATGAT 8782
 QY 1681 ATGTAAAGCCATTTTCTCAAGATGATGATACCTCTTAAGATTTGATGAATGCA 1740
 DB 8783 ATGTAAAGCCATTTTCTCAAGATGATGATACCTCTTAAGATTTGATGAATGCA 8842
 QY 1741 TTAACTTTTCAAGCTCTGAGTGTGATTTAGTGCATGAGGCAATTAATAGTCA 1800
 DB 8843 TTAACTTTTCAAGCTCTGAGTGTGATTTAGTGCATGAGGCAATTAATAGTCA 8902
 QY 1801 TGTGCAAAAGTATGACCTTAAATATTTGATGATGACCACTGCTCTTGG 1860
 DB 8903 TGTGCAAAAGTATGACCTTAAATATTTGATGATGACCACTGCTCTTGG 8962
 QY 1861 GAAAAAGTATGATTAATCTCTTGAAGTCTTTCCTCCCAAGTATGAGAA 1920
 DB 8963 GAAAAAGTATGATTAATCTCTTGAAGTCTTTCCTCCCAAGTATGAGAA 9022
 QY 1921 AATAATCTCTGCTGAGCTGAGAAACAGCTTCTGTTCTGCTGATATTTGTTAGT 1980
 DB 9023 AATAATCTCTGCTGAGCTGAGAAACAGCTTCTGTTCTGCTGATATTTGTTAGT 9082
 QY 1981 TTTTAATAGTCAATTTGATTAAGACCTTGTGGCTCCCAAGCTAAGTGAAGTTGATC 2040
 DB 9083 TTTTAATAGTCAATTTGATTAAGACCTTGTGGCTCCCAAGCTAAGTGAAGTTGATC 9142
 QY 2041 CCTACAGAGGCCATCTCAATTTAGAGAACAAAGCCCAATCTCTGCTCCAGACCTTA 2100
 DB 9143 CCTACAGAGGCCATCTCAATTTAGAGAACAAAGCCCAATCTCTGCTCCAGACCTTA 9202
 QY 2101 CCCCCAATCCCTGCGAGGTGTCTGCTCTGCTGCTCAATGAGAACTGGCAAGAAATAC 2160
 DB 9203 CCCCCAATCCCTGCGAGGTGTCTGCTCTGCTGCTCAATGAGAACTGGCAAGAAATAC 9262
 QY 2161 TAGAGGTGCGACAGTACTAGAGAAATGAGAAATGAGCACTACTGAGAGGAGA 2220
 DB 9263 TAGAGGTGCGACAGTACTAGAGAAATGAGAAATGAGCACTACTGAGAGGAGA 9322
 QY 2221 AAAATGGGCAACCTTCATGATGATGTTCTTTCCTTCTGTTGTTCAATGCTCGATA 2280
 DB 9323 AAAATGGGCAACCTTCATGATGATGTTCTTTCCTTCTGTTGTTCAATGCTCGATA 9382
 QY 2281 TAAATTTACAGAGGTGATATTAATCAATTTTCTTACCACTGAGAAAGCTGAGAAA 2340
 DB 9383 TAAATTTACAGAGGTGATATTAATCAATTTTCTTACCACTGAGAAAGCTGAGAAA 9442
 QY 2341 CTTGTTTACCATCATTAATTAATCTTCTTAAGTCAATTTCTAATTTCTAATGATCA 2400
 DB 9443 CTTGTTTACCATCATTAATTAATCTTCTTAAGTCAATTTCTAATTTCTAATGATCA 9502
 QY 2401 AATAGCTGACAAATATCTCTTGTATTAATTAATTAATTAATTAATTAATTAAT 2460
 DB 9503 AATAGCTGACAAATATCTCTTGTATTAATTAATTAATTAATTAATTAATTAAT 9562
 QY 2461 TAAATCTGCACTTTAGATGAGAAATTAATTAATTAATTAATTAATTAATTAAT 2520
 DB 9563 TAAATCTGCACTTTAGATGAGAAATTAATTAATTAATTAATTAATTAATTAAT 9622
 QY 2521 TTTTCATATGAAATTAATTAATTTTAAAGTATTAATTAATTAATTAATTAATTAAT 2580
 DB 9623 TTTTCATATGAAATTAATTAATTTTAAAGTATTAATTAATTAATTAATTAATTAAT 9682
 QY 2581 TAGATTTTCAATTAATGATTAAGATTAATTAATTAATTAATTAATTAATTAAT 2640
 DB 9683 TAGATTTTCAATTAATGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 9742
 QY 2641 GAAAAATCTTTTACATGATCAAGTCTTTTCAATTAATTAATTAATTAATTAATTAAT 2700
 DB 9743 GAAAAATCTTTTACATGATCAAGTCTTTTCAATTAATTAATTAATTAATTAATTAAT 9802

QY 2701 TTCTAATTAATACAGACTAATAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
 DB 9803 TTCTAATTAATACAGACTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9862
 QY 2761 ATGCTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2820
 DB 9863 ATGCTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9922
 QY 2821 TAAGTATGAAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2880
 DB 9923 TAAGTATGAAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9982
 QY 2881 ATAAAGTGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940
 DB 9983 ATAAAGTGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10042
 QY 2941 ATGTCAATACAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000
 DB 10043 ATGTCAATACAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10102
 QY 3001 TCAGCTCACCCCTGACCTGTAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
 DB 10103 TCAGCTCACCCCTGACCTGTAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 10162
 QY 3061 ACCTGTGACAGCGAGATTTATGAGAGCAAGCAATCAAGATCCGACGACACTTG 3120
 DB 10163 ACCTGTGACAGCGAGATTTATGAGAGCAAGCAATCAAGATCCGACGACACTTG 10222
 QY 3121 TCTCATCAAAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
 DB 10223 TCTCATCAAAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10282
 QY 3181 TGGCAAGTGTCTCTGACACTGAGGAGGCAATTAATTAATTAATTAATTAATTAATTAAT 3240
 DB 10283 TGGCAAGTGTCTCTGACACTGAGGAGGCAATTAATTAATTAATTAATTAATTAATTAAT 10342
 QY 3241 AAAAGAAAGTAAAGGAGAAAGTAAAGCAAGGAAAGATTTGATTTTAAACC 3300
 DB 10343 AAAAGAAAGTAAAGGAGAAAGTAAAGCAAGGAAAGATTTGATTTTAAACC 10402
 QY 3301 ATGCAAAAATTCGCAATCTCTGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 3360
 DB 10403 ATGCAAAAATTCGCAATCTCTGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 10462
 QY 3361 GTGATCTGAATGAGAAACAGGAGCAGAG 3420
 DB 10463 GTGATCTGAATGAGAAACAGGAGCAGAG 10492

RESULT 2
 AAC91773
 ID AAC91773 standard; DNA; 3438 BP.
 XX
 AC AAC91773;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Human myostatin gene promoter.
 XX
 XX Human myostatin gene promoter; growth/differentiation factor 8; GDF8;
 KW negative regulator; skeletal muscle growth; reporter plasmid;
 KW anti-myostatin antibody; expression inhibition; activity inhibition;
 KW drug screening; muscle wasting disorder; AIDS; multiple sclerosis;
 KW cancer; ageing; meat production; ds.
 XX
 OS Homo sapiens.
 XX
 XX W0200077206-A2.
 XX
 XX 21-DEC-2000.
 PD
 XX 09-JUN-2000; 2000WC-US015868.
 PF
 XX

PR 10-JUN-1999; 9905-00329685.

XX (ABBO) ABBOTT LAB.

PI Wu-Mong JR, Wang J,

DR WPI; 2001-071272/08.

XX Promoter which regulates expression of myostatin gene, whose inhibitors
PT are useful for preventing conditions involving muscle wasting such as
PT cancer, multiple sclerosis, aging and acquired immune deficiency
PT syndrome.

XX Claim 1; Fig 2; 31pp; English.

XX The invention relates to the human myostatin promoter. Myostatin (also
CC known as growth/differentiation factor 8 and GDF8) is a member of the TGF
CC -beta superfamily and is a negative regulator of skeletal muscle growth
CC and development. The invention also encompasses a reporter plasmid
CC comprising a reporter gene operably linked to the human myostatin
CC promoter, and host cells comprising the reporter plasmid. The invention
CC further relates to an antibody against human myostatin. The invention
CC screening for compounds which inhibit myostatin promoter activation,
CC which inhibit myostatin expression, or which interfere with the binding
CC of myostatin to a myostatin receptor. Compounds which inhibit myostatin
CC expression or function are useful for preventing disorders which involve
CC muscle wasting, such as AIDS, cancer, multiple sclerosis and aging. Anti-
CC myostatin antibodies may also be used in veterinary medicine to prevent
CC muscle wasting in ageing or diseased animals, and in agricultural
CC applications to increase meat production in livestock. The present
CC sequence represents the human myostatin promoter

XX Sequence 3438 BP; 1151 A; 628 C; 568 G; 1091 T; 0 U; 0 Other;

XX Query Match 43.1%; Score 1462.2; DB 4; Length 3438;

XX Best Local Similarity 72.7%; Pred. No. 5.2e-277;

XX Matches 2475; Conservative 0; Mismatches 743; Indels 187; Gaps 38;

QY 1 TGGTCTCCGAGCCATGAGAGGTTGCAATCCCGGACATGAGAGGAAACACGCTTC 60
DB 119 TTCTATTCTGGGGCTCAGAGAGGTTGCAATCACTGCGACGTGGAGAGGAAATTAACCTTA 178
QY 61 TTGAAATTTCTGAGTACCTTTAATTCATTCAATGCTGACCTCCGA-----GATTGG 114
DB 179 CTTAAATTTCTTCACTTCTTCTTCAACCATTCATACCTGTTCTCTAAGAGAGTTGATTAG 238
QY 115 ATGAGGTGACCTCATTTATTTACAGGTGTTATGTTTACCCCATCTGCGCAACA-TGA 173
DB 239 ATATGATGCAATTTCTCCCTATTTCAATGATGATGATGATGATGATGATGATGATGATG 298
QY 174 ATGAGGTGATCTGAGACACGCTGCCCCGAGGGAATCAAGAGTTGATCCCGTCAAGCT 233
DB 299 GTTAGCAAAATATGAACTAATGTTCCCAAGGGGAAAAACAGGTTAGTGAAGCCCTG--- 355
QY 234 CTGCTACGTTTTTTTGTAAACCAATCAATTAATACCTTTGTTGTGCAATTTCTGTTTT 293
DB 356 GTATATACATTTATCATCAACCAATCAATATATACCTTTGTTATGATGATGATGATGAT 415
QY 294 AAAATATCTTATTTATATAGTACTGCTAATTTCTTCAACATTTGTTGTCACACCAAAAG 353
DB 416 CAATTA-CATTATTTATATATATATTTGTCATTTCACTGAACCTGACACGCA-----G 470
QY 354 CCTTATTAATGAGCCCTGAATGAGCTTATACATACACATTTGTTTTTCTATGAGGA 413
DB 471 CAGACATATACACAGGCTGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 527
QY 414 AAAATTTTCTTCACTGCTGCGACAGCCTTTCTGCTTAAATTTGTTGCAAAATATATA 473
DB 528 AACTTTTCTTCTTCACTGCTGCGACATACAGCCTT-----TTAAATTTGTTGTTAAATATA 582
QY 474 TAACATGAAGCTGATATTTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 533
DB 583 ACATTTAGCA-----TCTTACCAATTTTGGTGTACAGTTCCGTTGCGCATTAAGCA 633

QY 534 CACTTACCTTGTGTGGAACTATGATCAATTCATCTTGAACCTTTTGTGATCTTCTCG 593
DB 634 CATTACACATGTTGTGTAACCATCAACACATTCATCTTGAGAAATTTTTCATCTTCCCA 693
QY 594 AATTGAAATCTTGATCTATTGACAGTATGCTTTCCCATATACACCGACCTTCTCTC 653
DB 694 GACTG--AACTGTATCTATCAACAGTAACTTACCCTCA-----GTGCTTACC 743
QY 654 CAGCCCCCGGACCAACATCTTCTTTCATGACATGATCCGTGCTGAGAAC--- 710
DB 744 CAGTCCCTGGCAACACATGCTAATCTTCCATGCGCTTTCTGTGATTTGGAACATAG 803
QY 711 -----CAGCCTTCAACACTACGCTTGGGGGCACTGT-----AAGCAACAGATC 756
DB 804 GAAACACGACAGCAGTACGATGATGCTTGGGGGCAATTTTAAAGCAAAATCAAT 863
QY 757 ACTTCCTACCGCCACCAATATGACACAAATATTAAGACATGTGGCATATGATGATTC 816
DB 864 AAGAGGACATTAATAAAGAACACAAATATATGTGAAACATGCGACTAAATAGACGGG 923
QY 817 AAAAAGGCTGTGCT--AGTATGAGGCTGAAC--AAGCAGAAATGACTAGT 871
DB 924 AAAAAGGAAATTTGTTTATATGATGAGCTGAACAAAGGCGGAACTTCTGTT 983
QY 872 TGAACCTGAGTGGGATCTGTGTGTGG--AAGCTCAATTTTTCATTTGCTGTGAT 929
DB 984 TGACCTCATCTGGGAACCTGTGCTGATGATGAGACTCAATTTTTCGCGCTGTGAAAT 1043
QY 930 ACGCAAAATGCTTATTAAGACCTGTAGGATGATGATGATGATGATGATGATGATGATG 989
DB 1044 GCGAGGAATGACTGAAAA--TGCTAGAAATATGATTTCAATAGATTAATTTTATG 1099
QY 990 CAAGACATAATATGACAGGAGTCCATGATTAACAGCATGACAT--GTGGAA 1047
DB 1100 AAGTAGCAAAATTTGTGTAATGACAAATATGATGATTAATGATTTGCTTACGGAAA 1159
QY 1048 TGATTAATCTTTTCTTCTTATTTCAAGCAGTAAAGAAAGCGC-----T 1093
DB 1160 TGACAACTTTTGTCACTTTTTCAGGTATTAAGATTAAGAGCTGGAACATGCTGTG 1219
QY 1094 CAGAGGCTGCTTACACCATTTTACTAGAGACTGACCTATGCTAGTGG----- 1144
DB 1220 AACCTGCTGCTTGAAGAAATTTTACTTGTGCTCAAGGCTTAAGAAATGAGCTTTGTGAC 1279
QY 1145 ---TAGCTGCAATTAACAACTGAAGCAATTTAGTTCAATGTGAGAGATTTTAAACA 1201
DB 1280 TGATTAATGCACTATGACCTTAAGCAATTTAGTTCAATGTGAGAGATTTTAAACA 1239
QY 1202 TAATCTCAACCCCTCTGATGAAACAGAGACTAAGTACTCAAGTACCAATTAACATCA 1261
DB 1340 TAATCTCAACCCCTCTGATGAAACAGAGACTAAGTACTCAAGTACTCAATTAACATTA 1399
QY 1262 CTTACTATATGACAGGACCTGATCTCAACATTTTACATGATTTTGAATTAACATGCCCC 1321
DB 1400 CTTACTATATGACAGGCTGATCTCAAGCAATTTTACATGATTTTAAATTAATATCCCC 1459
QY 1322 CACACCTATGAGAGAGCTGAAGTTAGAG-----AAGTATCTCAATCTTATTAACACA 1376
DB 1460 CAACACCTATGAGAGAGCTGAAGTTAGAGAGTTAAGTATCTCAATCTTATTAACACA 1519
QY 1377 GT-----GGCAACCTGATCTGAATCTGATCTATCACTCAAGACCTGAGATC 1428
DB 1520 GTTAAAGTGGCAAGGTGATGATTAACATCAAGCTGATCTCAAGACCTGAGATC 1579
QY 1429 CCAATGCTACACAAATTTTATCAAGTAAAGGAAAAAGATTTGATTTGCTCAAG 1488
DB 1580 TCAATTAACCTGCTATTAACAAATTTTAAAGCATATTAAGATTTTGAATTTTGAATTT 1639
QY 1489 TGTATAGGGGCAATATGATCA---ATTATTAACATTAAGATTTTATGTTGAAAAATA 1545
DB 1640 TGTATAGTGAATGATGATTAACATGATCAATTAAGATTTTGAAGATTTTCACTTGAATATG 1699

QY	1546	AATTATCAACAATATAAACTTTATTAAGCCGTGATCTTAATCTGCTCCGCAACAAAGCT	1605
Db	1700	AATTATCAACAATATAAACTTTATTAAGCCGTGATCTTAATCTGCTCCGCAACAAAGCT	1759
QY	1506	ATCTGAATACCTTCAGGGCATCTGATTTGCTGCTTTCTTAAATCTTAAATGATGGG	1665
Db	1760	ATCTGAATACCTTCAGGGCATCTGATTTGCTGCTTTCTTAAATCTTAAATGATGGG	1818
QY	1666	CAATCTTAATGATTTATGTAAGGCCA-TTTTCTCAAGAGATGTAATACCTCTTAAGA	1724
Db	1819	CAATCTTAATGATTTATGTAAGGCCA-TTTTCTCAAGAGATGTAATACCTCTTAAAGA	1878
QY	1725	ATTATGATGAATAATGATTAATCTTTCAAGGCTACTGATGTAATTTAGTGACGTGAAGCA	1784
Db	1879	ATTATGATGAATAATGATTAATCTTTCAAGGCTACTGATGTAATTTAGTGACGTGAAGCA	1938
QY	1842	ACCACTGCACTCTCTTGG-GGAAAAAGTAATGATTAATCTCTTTAGAGCTCTTAAGCT	1900
Db	1999	GCCACTGATATCTCTTGGAAAAAAAAGTAATGACATAATCTTAGGAATCCTTAAGCT	2058
QY	1901	TCCCAAAAGTAATGTAAGAAAAATTAATCTCTGCTGGCCGTGAAGAACGCTTCTGTTCTTG	1966
Db	2059	TCCCAAAAGTAATGTAAGAAAAATTAATCTCTGCTGGCCGTGAAGAACGCTTCTGTTCTTG	2118
QY	1961	CTGCTATATTTGTTTAGTTTAAATGATCAATTTGATAGACCTTGCTGCCAAG	2020
Db	2119	CTGCTATATTTGTTTAGTTTAAATGATCAATTTGATAGACCTTGCTGCCAAG	2178
QY	2021	CTAAGGTTGAGAGTTTGATCCCTTACAGAGGCCACTTCAATTTAGAGAACAAAAAGCCCA	2080
Db	2179	CTAAGGTTGAGAGTTTGATCCCTTACAGAGGCCACTTCAATTTAGAGAACAAAAAGCTCTA	2238
QY	2081	TTCTCTGTCCCAAGCCCTTACCCCAAAATCCCTGCAGAGTGTCTCCCTCGGTCAAAATGA	2140
Db	2239	TTCTCTGTCCCAAGCCCTTACCCCAAAATCCCTGCAGAGTGTCTCCCTCGGTCAAAATGA	2298
QY	2141	GAATCGGCAAGAAAGTACTAGAGAGTGCACAGTACTAGGAAGTAAAAATGAGACTA	2200
Db	2299	GAATCGGCAAGAAAGTACTAGAGAGTGCACAGTACTAGGAAGTAAAAATGAGACTA	2355
QY	2201	GCACACTACTAGAGAGCAAGAAAAATGGCACCTTCATGATGTTCTTCCCTTCT	2260
Db	2326	TCCGAGATATGGGAAACAAACAAAGGTCAACCCCTTAATCAGATGCTCTTCTCTTAT	2385
QY	2261	GTTTTCACAAATGCTCCCATATATTTACAGAGGGTAGTA-----ACTACATTTTT	2312
Db	2386	GTTTTCACAAATGCTCCCATATATTTACAGAGGGTAGTA-----ACTACATTTTT	2445
QY	2313	TCTTTTACCACTGGAAGCTGAGGAAAACTTTGTAACCATCATTAATTCACATATCTTC	2372
Db	2446	ATATTTTACTGCTGGAATCTTAGGCAACGTATATATCTCGCATGCCAGTTATATAA	2505
QY	2373	TAACTCATCTATGTTATTTCTAAG-----TCAATAGCTACAAATATCTCTTTGTATA	2426
Db	2506	TTCAATATCTTAGCTATGTTCAAGAGATTTTCTTAAGTGGCATTAACCTCTTTGTATA	2555
QY	2427	TAAACATGAAAAACACATCTCTGAGCAATTAATCTGCAACTTTAGATAGGAAGT	2486
Db	2566	TAAACATG-AAAAACACATCTCTGAGTATGTTAATCTCATCTTTAGATAGGA--	2622
QY	2487	AACCTTAATACATGCAATTTGAACGAAATCAATTTTCATATGTAATAAAGATAT-----	2542
Db	2623	---ATAATAGCACTCAGTCACAAAGTTTCAGTATATTTTCATATTAATAAGACATGA	2679
QY	2543	-TATTTAAAGTAATTCATGAGCAATTTAATATTA--AAGTAGATTTTCAATATGTT	2600
Db	2680	CTATGTAATAATATTCATGAGCAATATGTTATTAATAACATGACTTCAATATTTACT	2739
QY	2601	AAGATTTTATTCAGGAAACAAATTTCTCAATTAATAGACAGAAATCTTTTATTAATATC	2660

Dd		2740	AAGAAATTAGCA - GAAAACAAGTTCTCAAATTAATGAGAAGAAATTCACACTAGATC	2798
Oy		2661	ACAGCTCTTTCATTGAAGTC- TTCGGAATAAATCGTATTTCTAAATTATACAGACTA	2719
Dd		2799	ATAATCTTAACTTTAATTCAGGCCTCCATAATTTTTAAATTTCTTAATTCCTGGCACTA	2858
Oy		2720	AAAAATTAATTAATTAACAATAAATTAATTTT- TAOTCAANTGCTTACTTAATAGTA	2778
Dd		2859	AAAAATTAATTAATTAACAATAAATAAATTTCTTACTTCAAAATACCTTGCTTAACAATA	2918
Oy		2779	TAAATCATTTTATTTTCTGAGGGAAGAAGATATCACTTTTAAATGATGAAGTAAAT	2838
Dd		2919	TAAATCATTTTATTTTGTGAGGAAGTAATTTTC- ATATTTTAAATATGATGAATAAAT	2977
Oy		2839	TAAATCATTTTACTTAATTAATTAATTTTAAAGTTTCACATATAAAGTAATTAAGATC	2898
Dd		2978	TAAATGACTTATTTAAATTAACAATAAGATGTTGTGAGGAATTAAGTAATTAAGTAC	3037
Oy		2899	TAACTGATATGTTATTTGTTAATAAGTTTTTAATTTTGGCATGTCAATACAGCCTT	2958
Dd		3038	A-----GTTTATATTAATTTGCCAACATAGACTTTTGTTTTCAAAATGTCAAAATATCTTTT	3093
Oy		2959	ATTATTCATGATTTATTTCCTTTAAAGAAGTACAAATGAATCAGCTCACCTTGACTG	3018
Dd		3094	ATTATTTGTGATTTATTTCTTTTATGAGTAGTCAAATGAATCAGCTCACCTTGACTG	3153
Oy		3019	TAAACAAATACTGTTTGGTGACTTGTGACAGACAGGCTTTTAACTCTGACAGCGAATT	3078
Dd		3154	TAAACAAATACTGCTTGTGTGACTTGGACAGACAGGCTTTTAACTCTGACAGCGAATT	3213
Oy		3079	CATTGTGAGCAGAAGCCATCACAGATCCCAGACACATCTGTCTCATCAAAATTTGAAT	3138
Dd		3214	CATTGTGAGCAGAAGCCATCAATAGATCTGACAGACATCTGTCTCATCAAAATTTGAAT	3273
Oy		3139	ATAAAAAGCCACTGTGAATACAGATATAAAGATTCACTGNGTGGCAAGTTGTCTCAG	3198
Dd		3274	ATAAAAAGCCACTGTGAATACAGATATAAAGATTCACTGNGTGGCAAGTTGTCTCAG	3333
Oy		3199	ACTGGGACGACATTAAAGTTTGGCTTGGCGTTACTCAAAAGCAAAAGATAAAGGA	3258
Dd		3334	ACTGTACATGATTAATAATTTTGGCTTGGCATTTACTCAAAAGCAAAAGATAAAGGA	3393
Oy		3259	AGAACTAAGAACAAAGGAAAGATTGATTTTAAACCATG	3303
Dd		3394	AGAACTAAGAACAAAGGAAAGATTGATTTTAAACCATG	3438
<hr/>				
RESULT 3				
AAFS861	ID	AAFS861	standard; DNA; 2482 BP.	
AAFS8661;	AC			
XX	XX	27-APR-2001	(first entry)	
XX	DE	Murine myostatin gene promoter.		
KW	Mouse; myostatin; promoter; cytosolic; anti-HIV; neuroprotective;			
KM	immunomodulator; cancer; muscular dystrophy; spinal cord injury;			
KW	neurodegenerative disorder; congestive obstructive pulmonary disease;			
KV	COPD; amyotrophic lateral sclerosis; acquired immunodeficiency syndrome;			
XX	AIDS; cachexia; ageing; da.			
OS	Mus musculus.			
PN	EPI072680-AI.			
PD	31-JAN-2001.			
PF	27-JUL-2000; 2000EP-00306396.			
PR	30-JUL-1999; 99US-0146540P.			

XX (PFIZ) PFIZER PROD INC.

XX Findly RC;

XX WPI; 2001-170964/18.

XX Novel myostatin gene promoter for regulating expression of heterologous
PT genes in cells or animals, for identifying compounds that inhibit its
PT activity or expression, which are useful for treating muscle disorders.

XX Claim 1; Fig 1; 24pp; English.

XX The present sequence is an isolated myostatin gene promoter. Regulating
CC the transcription activities or expression of the promoter is useful for
CC treating muscle associated disorders such as cancer, muscular dystrophy,
CC spinal cord injury, neurodegenerative disorders, traumatic injury,
CC congestive obstructive pulmonary disease (COPD), amyotrophic lateral
CC sclerosis, acquired immunodeficiency syndrome (AIDS), cachexia or ageing.
CC The promoter is useful for modulating the expression of the myosin gene
CC for producing animal food products having increased muscle and protein
CC content and reduced fat and cholesterol content, for regulating
CC expression of a heterologous gene in cells or animals, for engineering host
CC cells, to screen for compounds that inactivate or inhibit its
CC transcription and expression, for inhibiting its expression in cells for
CC the promotion of muscle growth, for diagnostic evaluation, genetic
CC testing and prognosis of a disease or disorder associated with myostatin
CC expression

XX Sequence 2482 BP; 784 A; 519 C; 474 G; 705 T; 0 U; 0 Other;

XX Query Match 15.5%; Score 525; DB 4; Length 2482;

XX Best Local Similarity 61.5%; Pred. No. 3.7e-93;

XX Matches 1256; Conservative 0; Mismatches 665; Indels 122; Gaps 21;

QY 1188 GATGAATTTACCAATCTCAACCCCTCGCATGAAACAGACATCAATCACTCACTA 1247
DB 521 GAACACCTTGAGAAATTTGAAAGTAAAGAGCTTGAACAAAGAGAGCCCTTCTGCT 580
QY 1248 CCACTTACAGTCACTTACTATATG-----ACAGGCACTGACTCAACATTTTCAAT 1289
DB 581 TCAAGTATTTATTTCTATGAAAGGAGTACATTTAGTACTTTATTTGCTAAATTTAT 640
QY 1300 GTATTTATTTGATTTCAATGCCCCCAACTCTATGAGAGAGTGAAGTTAGAGATTC 1359
DB 641 GCCCTCAACCCCTTATTTAGTGAAGAACTAAAGATTAAGAGAGTCTGCGCTT 700
QY 1360 TCATTCATTTATTACACAGTGGCAAGTGAATCTGAGTCTAATCCAACTCCAGCA 1419
DB 701 TGTATGACATTTAAGAGAGGCAAAATGAGATTTGAACTCAGTTTATTTGACTTTCAG 760
QY 1420 CTTGAGATTCCTAATTTGCTACAAATTTCAATCAAGTTAAAGGAAAGATTTGATTT 1479
DB 761 TCTCAGCTCAAAATGGGAGTACAGTCTTAATAAATAAATAAATCAAGATCA-----ATT 816
QY 1480 GCTCAGAGTGTATAGGGGCAATATGTTA--AAATTAACATTAACAAGATTTATATGT 1536
DB 817 CCTCTGAGTATATAGAGCAATGTGTATATTTATTTGATCATGAAAGATTTATATGC 876
QY 1537 TGAATAATTAATTTATCAAAATTAATAAATTTATAGCTGATTAATCTGCTCCGCA 1586
DB 877 AGAAATAATGATTTTCCAGAAATTTGACTTTATAGGCTGCTCTAATATTTGCTTGTGA 936
QY 1597 ACAAGATATCTGAAATCTTTCAAGGAGCTGTTTGTGCTGTTTCTTATATCTTT 1656
DB 937 --TAAAGAGGCGCAGATCACTCAGGGTGTCTGTTGTGTCTGTTTCTTATCTTT 994
QY 1657 AATGATGGGCAATCTAATGATTTATGAGGCA-----TTTTTCTCAAGAGATGTAG 1711
DB 995 AATGATGGGCAATCTAATGATTTATGAGGCACTTTTCTTCTCAAGAGATGTAG 1054
QY 1712 ATACCTTTAAGATTTGATGAAATGATTAATTTTCAAGCTACGATGATTTTGA 1771
DB 1712 ATACCTTTAAGATTTGATGAAATGATTAATTTTCAAGCTACGATGATTTTGA 1771

DB 1055 ATGCCCTCTTAAAAATTTGATGAAATGATTAATTTTCAAGTACTGAGCTGATTTTA 1114
QY 1772 GTGCACTGAGGAGCAATTAATTT--AGTGTACATGTGCGAAAGTACCTTAATAATAA 1828
DB 1115 GTTCACTGAGGAGCAATTAATTTGATGATTTATCTGATCAGAAATGATGATGATTAATA 1174
QY 1829 ATATTTGATGATGAAACACTGCACTCTTGGGG-----AAAAAGTAATGATTAAC 1880
DB 1175 ATATTTATATCAAGCCACATAGCTCTTGGGGGTGTGTGTAAGGGGAGATTAATAA 1234
QY 1881 TCTCTTGAAGTCTTGTAGCTTCCCAAAAGATGAGAAATAATCTCTGAGGCTG 1940
DB 1235 ATTTAAAGATCTCTCAAGCTTCCAAACAGAGAGAA-----CTCTGAGGCTG 1285
QY 1941 GAACACCTTCTGTTCTTCTGCTGCTATTTTGTATTTGATTTTATATGATTTGATTT 2000
DB 1286 GAACGCTCTGCTGCTT--GCTGCTGTTTGTGTCAGCTCTTTAAGATTTACCCCAT 1343
QY 2001 AGACCTTGTGCTCCCAAGGCTAAAGTTGAGATTTGATCTCTAAGAGGCACTTCAAT 2060
DB 1344 CGATCTTGTGCTCTTAAAGCAGAGGTTGAGATTTGATCTCTGAGAGGCACTTAAAT 1403
QY 2061 TTAAAGAAACAAAGCCCAATCTCTGCTCCAGACCTTACCCCAATCTCTGCAAGT 2120
DB 1404 TCAGAGAAACAAAGACCAATCTCTGCT--CCTAGACTCTGAGCCCAAGATCTCTGCAAGT 1462
QY 2121 TCTGCTCTGCTGCTCAATGAGAACTGCAAGGAGTAACTAGAGAGTCTGCAAGTACTA 2180
DB 1463 TCTGCTCTGCTGCTCAATGAGAGCTGCAAGGAGTAACTAGAGAGTCTGCAAGTACTA 1518
QY 2181 GGAAGTAAAGAAATGAGCTAGACACATCTGAGAGAGAAATAATGGGCACTTCAATGA 2240
DB 1519 -----GGAGAGCAAAAGAGACACCTTCAATGA 1547
QY 2241 TGTGCTCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2300
DB 1548 ---TGCAGCTGCTTGT 1603
QY 2301 ACTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2351
DB 1604 ---TTGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1660
QY 2352 ATCATTAATTTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2411
DB 1661 TACTTACAAATTCATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1720
QY 2412 ATATCTCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2471
DB 1721 GATTTCTGTTGGGCAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1780
QY 2472 TTTAGATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2531
DB 1781 ATCTCTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1840
QY 2532 ATTAAGATTTAT--TTAAAGTATTTCTATGAGCAATTAATTAATTAATTAATTAATTA 2586
DB 1841 GTCAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1900
QY 2587 TTTCAATATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2646
DB 1901 TTGTAATATGTTCTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1960
QY 2647 CTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2706
DB 1961 TTTTAAAGAGTCAAAATCACTGCTTCTTAAATTAATTAATTAATTAATTAATTAATTAATTA 2020
QY 2707 TTAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2766
DB 2021 TTCT-----CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2068
QY 2767 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2826
DB 2069 GTTTTGTAAACAAAGCACTCAAGTCTCAAGATTAATTAATTAATTAATTAATTAATTAATTA 2127

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH mutation replace(284..289, agatct)
 FT /tag= b
 FT /note= "Mutation resulted in 25% decreased luciferase activity"
 FT misc_feature
 FT /tag= a
 FT /note= "Portion of promoter element which is claimed as SEQ ID NO:6"
 FT 393..398
 FT /tag= c
 FT /note= "Mutation resulted in increased luciferase activity"
 FT mutation
 FT /tag= d
 FT /note= "Mutation resulted in 40% decreased luciferase activity"
 XX WO200004051-A2.
 XX 27-JAN-2000.
 XX 15-JUL-1999; 99WO-US016026.
 XX 15-JUL-1998; 98US-0092865P.
 XX 08-MAR-1999; 99US-0123270P.
 XX (META-) METAMORPHIX INC.
 XX Liang L;
 XX WPI; 2000-182405/16.
 XX Novel promoter sequence for tissue specific gene expression in gene therapy, for treating muscle associated disorders, e.g. muscular dystrophy.
 PT Claim 4; Fig 2; 40pp; English.
 XX
 CC The patent discloses the use of growth differentiation factor (GDF) gene promoters for tissue specific gene expression in gene therapy. These elements are located upstream of the 5' end of a GDF gene. GDF promoters ensure proper binding and activation of RNA polymerase and influence transcription levels. They are used for muscle specific expression of a gene by transfecting it into a muscle cell and thus treatment of muscle associated disorders e.g. cancer, muscular dystrophy, spinal cord and traumatic injury, congestive obstructive pulmonary disease, AIDS, cachexia, and obesity or its related disorders (e.g. diabetes) and disorders related to abnormal proliferation of adipocytes by gene therapy. The present sequence is human GDF-8 promoter element used in the construction of luciferase reporter construct pGL-3-basle for modifying luciferase activity. This construct is used in screening for regulatory compounds of GDF expression
 CC
 CC Sequence 649 BP; 240 A; 89 C; 95 G; 225 T; 0 U; 0 Other;
 XX
 SQ
 Query Match 12.5%; Score 422.4; DB 3; Length 649;
 Best Local Similarity 82.0%; Fred. No. 3.7e-73;
 Matches 536; Conservative 0; Mismatches 111; Indels 7; Gaps 4;
 QY 2652 ACTAGTACAGCTTTTCATTAAAGTC-TTCGTAATAATCGTATTTCTAATAT 2710
 DB 1 ACTAGTACATTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCT 60
 QY 2711 ACAAGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2769
 DB 61 TTGGCACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCT 120
 QY 2770 TAAATAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 2829
 DB 121 TAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGT 179

QY 2830 AGGTAAATTAAGATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 2889
 DB 180 AGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 239
 QY 2890 AATTAAGTCAAGTATATGTTATGTTAATTAAGTTTAAATTTTTCGATGTCAT 2949
 DB 240 TTTAAG----TACGTTTATTAATTAATTAATTAATTAATTAATTAATTAATGA 295
 QY 2950 ACAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 3009
 DB 296 ATATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 355
 QY 3010 CTTGACTTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 3069
 DB 356 CTTGACTTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 415
 QY 3070 AGCGAGTTCATTGTGAGCAAGGCCAATCAAGATCCGACGACACTTGTCTCATCA 3129
 DB 416 AGCGAGTTCATTGTGAGCAAGGCCAATCAAGATCCGACGACACTTGTCTCATCTA 475
 QY 3130 AGTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 3189
 DB 476 AGTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGA 535
 QY 3190 GTCTCTGAGCTGGGCAAGGCAATTAAGCTTGGCTTGGCTTAAAGCAAAAGAAA 3249
 DB 536 GTCTCTGAGCTGGGCAAGGCAATTAAGCTTGGCTTGGCTTAAAGCAAAAGAAA 595
 QY 3250 GTAAAGGAGAGTAATTAAGCAAGGCAAAAGATTTGATTTTAAACCATG 3303
 DB 596 GTAAAGGAGAGTAATTAAGCAAGGCAAAAGATTTGATTTTAAATCATG 649

RESULT 7

AFS8662/c

ID AFS8662 standard; DNA; 1537 BP.

XX AFS8662;

DT 27-APR-2001 (first entry)

XX Porcine myostatin gene promoter.

KW Porcine; pig; myostatin; promoter; cytoskeletal; anti-HIV; neuroprotective;

KW immunomodulator; cancer; muscular dystrophy; spinal cord injury;

KW neurodegenerative disorder; congestive obstructive pulmonary disease;

KW COPD; amyotrophic lateral sclerosis; acquired immunodeficiency syndrome;

KW AIDS; cachexia; ageing; de.

XX Sus scrofa.

XX EPI072680-A1.

XX 31-JAN-2001.

XX 27-JUL-2000; 2000EP-00306396.

XX 30-JUL-1999; 99US-0146540P.

XX (PRIZ) PFIZER PROD INC.

XX Findly RC;

XX WPI; 2001-170964/18.

XX Novel myostatin gene promoter for regulating expression of heterologous

XX genes in cells or animals, for identifying compounds that inhibit its

XX activity or expression, which are useful for treating muscle disorders.

XX Disclosure; Page 12-13; 24pp; English.

XX The present sequence was used for comparison studies with a myostatin

CC gene promoter sequence isolated from Mus musculus. Regulating the
 CC transcription activities or expression of the isolated promoter is useful
 CC for treating muscle associated disorders such as cancer, muscular
 CC dystrophy, spinal cord injury, neurodegenerative disorders, traumatic
 CC injury, congestive obstructive pulmonary disease (COPD), amyotrophic
 CC lateral sclerosis, acquired immunodeficiency syndrome (AIDS), cachexia or
 CC ageing. The promoter is useful for modulating the expression of the
 CC myosin gene for producing animal food products having increased muscle
 CC and protein content and reduced fat and cholesterol content, for
 CC regulating expression of a heterologous gene in cells or animals, to
 CC engineer host cells, to screen for compounds that inactivate or inhibit
 CC its transcription and expression, for inhibiting its expression in cells
 CC for the promotion of muscle growth, for diagnostic evaluation, genetic
 CC testing and prognosis of a disease or disorder associated with myostatin
 CC expression
 CC
 XX Sequence 1537 BP; 519 A; 259 C; 236 G; 523 T; 0 U; 0 Other;
 XX
 XX
 Query March 9.8%; Score 332.8; DB 4; Length 1537;
 Best Local Similarity 63.9%; Pred. No. 1.7e-55;
 Matches 555; Conservative 0; Mismatches 302; Indels 11; Gaps 3;
 Oy 2485 GTAACCTTAATCTAGTCAATGAACTGAAATACAAATTTGATGATGAAATGTA 2544
 Db 1223 GAAACCTTAATGATGATTAATTAATGATATTTTACACTGATCTTAATAAGTTG 1164
 Oy 2545 TTTAAAGTAATTCAGAGCAATTTAATTAAGTAGATTTTCATTATGTGTAGA 2604
 Db 1163 AAATGCTTTTCTCAAAAAAATTAATGATTTTACATTTTAAAGAAACATTTGGG 1104
 Oy 2605 ATTATTCAGGGAACAAGTTCTCAATTTAGAGAAAATCTTTACTAGTACAG 2664
 Db 1103 AAAAAAGATTTTATTTGCTGTTTAAATTTTATTTAGCTGCTGATACAGAAAAATC 1044
 Oy 2665 TCTTTCAATTAAGTCTCTGAAATAATCTGATTTTCTAATTAATACAGATAAAT 2724
 Db 1043 AGATTTAAATAGGAAGACCTGAATTA---AATTTAAGATGTGATACAGTAAATA 988
 Oy 2725 AATTTAATTAACAATAAATTAATTTTACTTCAATAGCTTACTTAATAGTAAAT 2784
 Db 987 TATTCAGCTAATTAATTTGAGAAACCTGTTTCTGACTAATTTCTAGTACATTAAGAAA 928
 Oy 2785 CATTTATTTCTGAGGAGAAAGCATATCACTTTTAAGTGAAGTGAATTAAGAT 2844
 Db 927 TCAATCTAATCTGTTAATTAATACAGAGAAATTTTAATTAATCTTTTATTCAG 868
 Oy 2845 TTATTCACCTTAATTAATTAATTTTAAAGTTTCACATATTAAGATTAAGTCTAAGT 2904
 Db 867 TGAATAATGCTGCTAGTTTCAACGATTTTCTCATCTTAAGTGAAGTAAATAT 808
 Oy 2905 TATATGTTATGTTAATAAAGTTTAAATTTTGGAGTCAATACAGCCTTATATT 2964
 Db 807 GCTCAGAAAGATGTGTTTTCATTTTATTAATCAAGAGATTAATGTCAGTTTCAATTAT 748
 Oy 2965 CATATATTTATCTTTTAAAGAGATTAAGTCAATGATGAGCTCACCTTGAATGAACA 3024
 Db 747 TGTAAATTAATTTCTTTTAAAGTAGTAAGTAAGTAATGCTCACCTTGAATGAACA 688
 Oy 3025 AATATCTGTTGCTGCTGTCAGACAGAGGTTTAACTCTGACAGCAGATTCATTGT 3084
 Db 687 AATATCTGTTGCTGCTGTCAGACAGAGGTTTAACTCTGACAGCAGATTCATTGT 628
 Oy 3085 GGAGCAAGAGCAATACAGATCCGAGAGCACTGTCTCATCAAAAGTTGAATTAATA 3144
 Db 627 GGAGCAAGAGCAATACAGATCCGAGAGCACTGTCTCATCAAAAGTTGAATTAATA 570
 Oy 3145 AGCCACTTGAATACAGTATAAAGATTGACTGCTGAGCAAGTGTCTCTCAGACTGGG 3204
 Db 569 ACCCACTTGAATACAGTATAAAGATTGACTGCTGAGCAAGTGTCTCTCAGACTGGG 510
 Oy 3205 CAGGCAATTAACGTTGGCTTGGCTTACTCAAAAGCAAAAGAAAGTAAGAGAAAGT 3264
 Db 509 CAGGCAATTAATTTTGGCTTGGCTTACTCAAAAGC-----AAAAGTAAGAGAAAGT 455

Oy 3265 AAGACAGAGGAAAGATTGTATGATTTTAAACCATGCAAAATCTGCTTT 3324
 Db 454 AAGACAGAGGAAAGATTGTATGATTTTAAATCATGCAAAATCTATGTT 3395
 Oy 3325 TATATTTACTTATTTATGCTGATTTG 3352
 Db 394 TATATTTACTGTTTATGCTGATCTG 367
 RESULT 8
 ID AAZ50113 standard; DNA; 396 BP.
 AC AAZ50113;
 XX
 XX 04-MAY-2000 (first entry)
 XX
 XX Mouse growth differentiation factor-8 promoter element.
 DE
 XX Growth differentiation factor gene promoter; GDF; gene therapy;
 KM tissue specific gene expression; cancer; AIDS; cachexia; mouse;
 KM muscle associated disorder; muscular dystrophy; spinal cord injury;
 KM traumatic injury; congestive obstructive pulmonary disease; obesity;
 KM diabetes; cyostatic; immunomodulator; antidiabetic; anorectic; cardiant;
 KM ds.
 OS Mus musculus.
 XX
 XX Key Location/Qualifiers
 FH 83..240
 FT misc_feature
 FT /tag= a
 FT /note= "Portion of GDF-8 promoter element claimed as SEQ
 ID NO.7"
 FT 91..96
 FT misc_binding
 FT /tag= c
 FT /bound_molety= "Transcription factor"
 FT /note= "Functions as a regulatory region"
 FT 145..150
 FT misc_binding
 FT /tag= d
 FT /bound_molety= "Transcription factor"
 FT /note= "Functions as a regulatory region"
 FT 165..170
 FT misc_binding
 FT /tag= e
 FT /bound_molety= "Transcription factor"
 FT /note= "Functions as a regulatory region"
 FT 393..236
 FT /tag= b
 PD WO200004051-A2.
 PD 27-JAN-2000.
 PF 15-JUL-1999; 99WO-US016026.
 XX
 XX 15-JUL-1998; 98US-0092865P.
 PR 08-MAR-1999; 99US-0123270P.
 XX
 XX (META-) METAMORPHIX INC.
 PA
 PI Liang L;
 DR WPI; 2000-182405/16.
 XX
 XX Novel promoter sequence for tissue specific gene expression in gene
 PT therapy, for treating muscle associated disorders, e.g. muscular
 PT dystrophy.
 XX
 XX Claim 5; Fig 7; 40pp; English.
 XX
 CC The patent discloses the use of growth differentiation factor (GDF) gene
 CC promoters for tissue specific gene expression in gene therapy. These
 CC elements are located upstream of the 5'end of a GDF gene. GDF promoters

CC ensure proper binding and activation of RNA polymerase and influence
 CC transcription levels. They are used for muscle specific expression of a
 CC gene by transfecting it into a muscle cell and thus treatment of muscle
 CC associated disorders e.g. cancer, muscular dystrophy, spinal cord and
 CC traumatic injury, congestive obstructive pulmonary disease, AIDS,
 CC cachexia, and obesity or its related disorders (e.g. diabetes) and
 CC disorders related to abnormal proliferation of adipocytes by gene
 CC therapy. The present sequence is mouse GDF-8 promoter element isolated by
 CC screening a mouse GDF-8 genomic clone. This sequence shows a high level
 CC of homology with human, pig and chicken GDF-8 promoter elements
 XX

SO Sequence 396 BP; 131 A; 70 C; 80 G; 115 T; 0 U; 0 Other;

Query Match 7.7%; Score 259.4; DB 3; Length 396;
 Best Local Similarity 82.3%; Pred. No. 3.1e-41;
 Matches 312; Conservative 0; Mismatches 61; Indels 6; Gaps 1;

Oy 2920 ATTAAGTTTATTTATTTTGGCATGTCACATACAGCTTATTTATTCATATGATTATTCCT 2979
 Db 18 ACACAGACTTCATATTCATCAATGTCACATATATCTTCATATGATTTGGGATTTATTTCA 77
 Oy 2980 TTTAAGATGATCAATGAATCAGTCACTCCTGACCTGTAACTAATATCTTTGGTGA 3039
 Db 78 TTTAAGATGATCAATGAATCAGTCACTCCTGACCTGTAACTAATATCTTTGGTGA 137
 Oy 3040 CTTGTGACAGACAGGGTTTAACTCTGACAGAGATTCATTTGTGACAGACCAAT 3099
 Db 138 CTTGTGACAGACAGGGTTTAACTCTGACAGAGATTCATTTGTGACAGACCAAT 197
 Oy 3100 CACAGATCCGACGACACTTGTCTCATCAAAAGTTGAATATTAAGCCACTTGGAAATAC 3159
 Db 198 CATGATCTTCAACGACACTTGTCTCATCAAAAGTTGAATATTAAGCCACTTGGAAATAC 257
 Oy 3160 AGTATTAAGATTCCTGTGTGCAAGTGTCTCTGACAGCTGGGACGATTAACGTT 3219
 Db 258 AGTATTAAGATTCCTGTGTGCAAGTGTCTCTGACAGCTGGGACGATTAACGTT 317
 Oy 3220 GCGTTGGCTTACTCAAAAGCAAAAGAAAGTAAAGAGAGATTAAGCAAGGAAAA 3279
 Db 318 CACTTGGCATTAATCAAAAGCAAAAGAAAGTAAAGAGAGATTAAGCAAGGAAAA 371
 Oy 3280 GATTGTATTTATTTTAAAA 3298
 Db 372 GATTGTCTGATTTTAAAA 390

RESULT 9
 ADH11572
 ID ADH11572 standard; cDNA; 2823 BP.

XX ADH11572;
 AC
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human bone morphogenic protein (BMP) polynucleotide #21.
 XX
 KW Human; bone morphogenic protein; BMP; gene; ss; weight gain;
 KW appetite suppression; fat mass reduction; cell sensitivity;
 KW glucose uptake; diabetes; insulin resistance; hyperglycemia;
 KW hypertension; coronary artery disease; renal failure; neuropathy;
 KW metabolic disorder; glucose metabolism disorder; endocrine disorder;
 KW obesity; weight loss; liver disorder; cartilage growth disorder;
 KW bone growth disorder; inflammation; aberrant cell growth; liver cancer.
 XX
 OS Homo sapiens.
 XX
 PN US200324501-A1.
 XX
 PD 04-DEC-2003.
 XX
 PF 14-FEB-2003; 2003US-00366345.
 XX
 PR 17-MAR-2000; 2000US-0190067P.
 XX
 PR

PR 16-MAR-2001; 2001US-00809269.
 PR 23-MAR-2001; 2001WO-US009229.
 PR 17-JAN-2002; 2002US-0348621P.
 PR 22-JAN-2002; 2002US-0349356P.
 PR 28-JAN-2002; 2002US-0351520P.
 PR 06-FEB-2002; 2002US-0354265P.
 PR 15-FEB-2002; 2002US-0356749P.
 PR 16-JAN-2003; 2003US-00345236.
 PA (YOUNG) YOUNG P. E.
 PA (RUBEN) RUBEN S. M.
 XX
 PI Young PE, Ruben SM;
 XX
 DR WPI: 2004-022075/02.
 DR P-PBDB; ADH11595.
 XX

PT New bone morphogenic protein polypeptides and polynucleotides, useful for
 PT diagnosing, preventing, treating or ameliorating a medical condition,
 PT e.g. diabetes, dyslipidemia, hypertension, coronary artery disease or
 PT neuropathy.

Example 1; SEQ ID NO 27; 224bp; English.

CC The invention relates to human bone morphogenic protein (BMP)
 CC polypeptides and the polynucleotides encoding them. The invention also
 CC relates to a method for limiting weight gain, suppressing appetite or
 CC reducing fat mass, comprising administering to a mammalian subject a
 CC therapeutic amount of a BMP polypeptide, and a method for increasing the
 CC sensitivity of a cell to insulin or increasing glucose uptake by a cell,
 CC comprising contacting the cell with a BMP polypeptide. The BMP
 CC polypeptides and polynucleotides are useful for diagnosing a pathological
 CC condition or a susceptibility to a pathological condition in a subject or
 CC for preventing, treating or ameliorating a medical condition, e.g.
 CC diabetes, insulin resistance, hyperglycemia, hypertension, coronary
 CC artery disease, renal failure, neuropathy, metabolic disorders, glucose
 CC metabolism disorder, endocrine disorders, obesity, weight loss, liver
 CC disorders, cartilage and bone growth disorders, inflammation or aberrant
 CC cell growth such as liver cancer. The BMP polypeptides and
 CC polynucleotides are also useful for regulating nutritional partitioning,
 CC limiting weight gain, suppressing appetite, reducing fat mass, increasing
 CC the sensitivity of a cell to insulin or increasing glucose uptake by a
 CC cell. This sequence represents a human BMP polynucleotide of the
 CC invention.
 XX

SO Sequence 2823 BP; 978 A; 457 C; 497 G; 891 T; 0 U; 0 Other;

Query Match 5.6%; Score 190; DB 12; Length 2823;
 Best Local Similarity 91.0%; Pred. No. 2e-27;
 Matches 202; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Oy 3168 AGATTCACCTGTGTGGCAAGTGTCTCTCAACATCGGCGAGCATTAACGTTGGCTGGC 3227
 Db 1 AGATTCACCTGTGTGGCAAGTGTCTCTCAACATCGTACATTAATAATTTCTGGC 60
 Oy 3228 GTTACTCAAAAGCAAAAGAAAGTAAAGAGAGTAAAGCAAGGAAAAAGATTGAT 3287
 Db 61 ATTACTCAAAAGCAAAAGAAAGTAAAGAGAGTAAAGCAAGGAAAAAGATTAT 120
 Oy 3288 TGATTTTAAAAACATGCAAAAGTCAAAATCTCTGTTTATTTTAACTATTATGCTGAT 3347
 Db 121 TGATTTTAAAAATCATGCAAAAGTCAAACTCTGTTTATTTTAACTGTTATGCTGAT 180
 Oy 3348 TGTTCCTGCGCCCATGTGATCTGAATGAGACAGGAGCAGAA 3389
 Db 181 TGTTCCTGCTCAAGTGAATTAATGAGAACAGTGAGCAGAAA 222

RESULT 10
 AAF58663
 ID AAF58663 standard; DNA; 200 BP.
 XX
 AC AAF58663;
 AC

XX 27-APR-2001 (first entry)

DE Murine myostatin gene promoter fragment.

XX Mouse; myostatin; promoter; cytostatic; anti-HIV; neuroprotective;
 XX immunomodulator; cancer; muscular dystrophy; spinal cord injury;
 KW neurodegenerative disorder; congestive obstructive pulmonary disease;
 KW CODD; amyotrophic lateral sclerosis; acquired immunodeficiency syndrome;
 KW AIDS; cachexia; ageing; ds.

OS Mus musculus.

XX EP1072680-A1.

XX 31-JAN-2001.

XX 27-JUL-2000; 2000EP-00306396.

XX 30-JUL-1999; 99US-0146540P.

XX (PRIZ) PRIZER PROD INC.

XX Findly RC;

XX WPI; 2001-170964/18.

XX Novel myostatin gene promoter for regulating expression of heterologous
 PT genes in cells or animals, for identifying compounds that inhibit its
 PT activity or expression, which are useful for treating muscle disorders.

PS Claim 3; Page 13; 24pp; English.

XX The present sequence is part of an isolated myostatin gene promoter.
 CC Regulating the transcription activities or expression of the promoter is
 CC useful for treating muscle associated disorders such as cancer, muscular
 CC dystrophy, spinal cord injury, neurodegenerative disorders, traumatic
 CC injury, congestive obstructive pulmonary disease (COPD), amyotrophic
 CC lateral sclerosis, acquired immunodeficiency syndrome (AIDS), cachexia or
 CC ageing. The promoter is useful for modulating the expression of the
 CC myosin gene for producing animal food products having increased muscle
 CC and protein content and reduced fat and cholesterol content, for
 CC regulating expression of a heterologous gene in cells or animals, to
 CC engineer host cells, to screen for compounds that inactivate or inhibit
 CC its transcription and expression, for inhibiting its expression in cells
 CC for the promotion of muscle growth, for diagnostic evaluation, genetic
 CC testing and prognosis of a disease or disorder associated with myostatin
 CC expression

XX Sequence 200 BP; 57 A; 44 C; 48 G; 51 T; 0 U; 0 Other;

XX Query Match 5.1%; Score 174.4; DB 4; Length 200;

XX Best Local Similarity 92.0%; Pred. No. 1.3e-24;

XX Matches 184; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2990 AGTCAATGATCAGCTCACCCTTGACTGTAAACAAATCTGTTGGTGAATGACAG 3049

DB 1 AGTCAATGATCAGCTCACCCTTGACTGTAAACAAATCTGTTGGTGAATGACAG 60

QY 3050 ACAGGCTTTAACTCTGACAGCAGGATTCATTGTGAGCAAGCAATCAAGATCC 3109

DB 61 ACAGGCTTTAACTCTGACAGCAGGATTCATTGTGAGCAAGCAATCAAGATCC 120

QY 3110 GACGACACTTGTCTCATCAAGTTGGAATTAATAAGCCACTTGGAAATCACTATTAAG 3169

DB 121 GACGACACTTGTCTCTCTTAAGTTGAATTAATAAGCCACTTGGAAATCACTATTAAG 180

QY 3170 ATTCACTGTTGGCAAGTT 3189

DB 181 ACTCCCTGGCGTGGCAGGTT 200

RESULT 11

AAZ50115
 ID AAZ50115 standard; DNA; 156 BP.

XX AAZ50115;

XX 04-MAY-2000 (first entry)

XX Pig growth differentiation factor-8 promoter element.

XX Growth differentiation factor gene promoter; GDF; gene therapy;
 KW tissue specific gene expression; cancer; AIDS; cachexia; mouse;
 KW muscle associated disorder; muscular dystrophy; spinal cord injury;
 KW traumatic injury; congestive obstructive pulmonary disease; obesity;
 KW diabetes; cytostatic; immunomodulator; antidiabetic; anorectic; cardiant;
 KW ds.

XX Sus scrofa.

XX Key Location/Qualifiers

XX misc_binding 9..14
 FT /tag= a
 FT /bound_molecy= "Transcription factor"

FT /note= "Functions as a regulatory region"

FT misc_binding 63..68
 FT /tag= b
 FT /bound_molecy= "Transcription factor"

FT /note= "Functions as a regulatory region"

FT misc_binding 83..88
 FT /tag= c
 FT /bound_molecy= "Transcription factor"

FT /note= "Functions as a regulatory region"

FT TATA_signal 152..156
 FT /tag= d

XX WO200004051-A2.

XX 27-JAN-2000.

XX 15-JUL-1999; 99WO-US016026.

XX 15-JUL-1998; 98US-0092865P.

XX 08-MAR-1999; 99US-0132270P.

XX (META-) METAMORPHIX INC.

XX Liang L;

XX WPI; 2000-182405/16.

XX Novel promoter sequence for tissue specific gene expression in gene
 PT therapy, for treating muscle associated disorders, e.g. muscular
 PT dystrophy.

PS Claim 5; Fig 8; 40pp; English.

XX The patent discloses the use of growth differentiation factor (GDF) gene
 CC promoters for tissue specific gene expression in gene therapy. These
 CC elements are located upstream of the 5' end of a GDF gene. GDF promoters
 CC ensure proper binding and activation of RNA polymerase and influence
 CC transcription levels. They are used for muscle specific expression of a
 CC gene by transfecting it into a muscle cell and thus treatment of muscle
 CC associated disorders e.g. cancer, muscular dystrophy, spinal cord and
 CC traumatic injury, congestive obstructive pulmonary disease, AIDS,
 CC cachexia, and obesity or its related disorders (e.g. diabetes) and
 CC disorders related to abnormal proliferation of adipocytes by gene
 CC therapy. The present sequence is pig GDF-8 promoter element isolated by
 CC screening a pig GDF-8 genomic clone obtained from GenBank Accession
 CC numbers AJ133580 and AF093798. This sequence shows a high level of
 CC homology with human, mouse and chicken GDF-8 promoter elements

XX Sequence 156 BP; 50 A; 31 C; 34 G; 41 T; 0 U; 0 Other;

XX Query Match 4.2%; Score 141.8; DB 3; Length 156;

[illegible]

```

RESULT 12
AAV89102
ID AAV89102 standard; cDNA; 336 BP.

```

AC AAV89102;
xy

DT 15-FEB-1999 (first entry)
 VV

DE EST clone BV995.
XX

Human; secreted protein; expressed sequence tag; EST; haematopoiesis; tissue growth; activin; inhibitor; chemotaxis; chemokines; haematopoietic

receptor; ligand; chromolytic; anti-inflammatory; cadherin; anti-tumour;
gene therapy; ss.

Homo sapiens.

PN W09845436-A2.

PD 15-OCT-1998.

PF 10-APR-1998; 98WO-US006955.

PR 10-APR-1997; 97US-00838821.

PA (GEMY) GENETICS INST INC.

PI Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;

XX
XX
WPT 1000 070077/AC

New	ncj	mucl	coc	i	d	q	c	t	e
XX									
PT									

PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT nutritive retina and colon, CNNA libraries

PS Claim 1: Page 111: 618nn: English
xx

The present sequence represents a human expressed sequence tag (EST). The polynucleotide, which is a secreted EST, and the encoded protein are predicted to have useful biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional activity, immune stimulating or suppressing activity, hematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, tumour inhibition activity. The polynucleotide may also be useful for gene therapy.

SQ Sequence 336 BP; 136 A; 52 C; 65 G; 83 T; 0 U; 0 Other;

Query Match 4.0%; Score 136.2; DB 2; Length 336;

Matches 150; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY	Db	QY	Db	QY	Db	QY	Db	QY	Db		
3217	TTTGCGCTTGCGGTACGCAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	3276	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	3325	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	3384	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	3433	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	3492	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA
4	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	63	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	122	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	181	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA	240	TTTGCGCGCTTGCGGTACGCTTAAAAAGCAAAAGAAATTAAGAGAAAGATTAACAACAAGGA		
3277	AAAGATGTATGTATGATTTTAAAAACCATGCAAAAACTGCAAAATCTGTGTTATATTAACCTA	3336	AAAGATGTATGTATGATTTTAAAAACCATGCAAAAACTGCAAAATCTGTGTTATATTAACCTA	3395	AAAGATGTATGTATGATTTTAAAAACCATGCAAAAACTGCAAAATCTGTGTTATATTAACCTA	3454	AAAGATGTATGTATGATTTTAAAAACCATGCAAAAACTGCAAAATCTGTGTTATATTAACCTA	3513	AAAGATGTATGTATGATTTTAAAAACCATGCAAAAACTGCAAAATCTGTGTTATATTAACCTA		
64	AAAGATGTATGTATGATTTTAAAAATATGCAAAAACTGCAACTCTGTGTTATATTAACCTG	123	AAAGATGTATGTATGATTTTAAAAATATGCAAAAACTGCAACTCTGTGTTATATTAACCTG	182	AAAGATGTATGTATGATTTTAAAAATATGCAAAAACTGCAACTCTGTGTTATATTAACCTG	241	AAAGATGTATGTATGATTTTAAAAATATGCAAAAACTGCAACTCTGTGTTATATTAACCTG	300	AAAGATGTATGTATGATTTTAAAAATATGCAAAAACTGCAACTCTGTGTTATATTAACCTG		
3337	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTGTAATGAGAAACAGGACGAGAA	3389	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTGTAATGAGAAACAGGACGAGAA	3440	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTGTAATGAGAAACAGGACGAGAA	3491	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTGTAATGAGAAACAGGACGAGAA	3542	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTGTAATGAGAAACAGGACGAGAA		
124	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTTAAATGAGAAACAGGACGAGAA	176	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTTAAATGAGAAACAGGACGAGAA	228	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTTAAATGAGAAACAGGACGAGAA	280	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTTAAATGAGAAACAGGACGAGAA	332	TTTATGCTGATTTGTTGCTGCGCCAGTGGATCTTAAATGAGAAACAGGACGAGAA		

RESULT 13

ID	Accession	Source	Size (bp)	Accession	Source	Size (bp)
AA024415	standard; cDNA	1195	BP.			

AC AAX24415;
VY

DT 07-JUN-1999 (first entry)
 YX

DE Bovine myostatin cDNA. XX

double muscling; muscle hypertrophy; transgenic animal; ss. myostatin; cattle; bovine; transdominant growth factor beta, 10

05 *Bos taurus*.

PH Key

ET : 330

13

XX

XX

10
11
12
13
14
15
16

XX 14-TH 1007. 07110 00001700

PR 15-JAN-1998; 98US-00007761.
YY

PA (UyLI-) UNIV LIEGE.
XX

PI Grobet L, Georges M, Poncelet D,
XX

DR WP1; 1999-120869/10
DR P-PSDB: AAW97884.

PT Increasing muscle mass in mammals - by decreasing myostatin expression

PS Claim 7; Page 54; 75pp; English.

CC This is nucleotide sequence of a cDNA cloneencoding bovine myostatin
CC (see AA0937884), a member of the transforming growth factor beta
CC superfamily. A mutant gene (see A0224416) in which the coding sequence
CC lacks an 11 bp consecutive sequence (see A0224421) has also been
CC sequenced. Cattle of the Belgian Blue breed homozygous for the mutant
CC gene are double-muscled. A new method of increasing muscle mass of a
CC mammal having myostatin-expressing muscle cells, comprises administration
CC of a nucleic acid molecule substantially complementary to at least a
CC portion of mRNA encoding myostatin and of sufficient length to reduce
CC myostatin expression and thus increase muscle mass. A ribozyme may also
CC be used. Also claimed are: a method for determining muscular hypertrophy
CC (MH) in a mammal using primers based upstream and downstream of the
CC mutation; a diagnostic kit for determining the genotype of a sample of
CC genetic material; a method for determining MH in a mammal; a method for
CC determining double muscling in a bovine animal; a method for determining
CC the myostatin genotype of an animal; purified myostatin; isolated nucleic
CC acids; a microbial host cell; a probe based on the myostatin gene
CC mutation; transgenic mammals having MH phenotype; a myostatin knockout

CC animal; and a transgenic bovine having a gene encoding active myostatin,
CC and a heterologous nucleotide sequence antisense to that gene, and
CC optionally further containing a gene encoding a nucleic acid sequence
CC with ribozyme activity in transcriptional association with the antisense
CC sequence. Primers are preferably based on genomic bovine myostatin DNA
CC (see AAX24464) and human myostatin cDNA (see AAX24418)
XX

SO Sequence 1195 BP; 379 A; 250 C; 261 G; 305 T; 0 U; 0 Other;

Query Match 3.8%; Score 127.6; DB 2; Length 1195;
Best Local Similarity 97.0%; Pred. No. 2.9e-15;
Matches 130; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3257 GAAGAGTAAGAAAGAGGAAAGATTTGATTTAAACATGCAAAACTGCAAA 3316
DB 2 GAAAGAAATTAAGAACGAGGAAAGATTTGATTTAAACATGCAAAACTGCAAA 61

QY 3317 TCTGTTTATTTATTTACTTATTTATGCTGATTTGTCGCCCATGATCTGATGAGA 3376
DB 62 TCTGTTTATTTATTTACTTATTTATGCTGATTTGTCGCCCATGATCTGATGAGA 121

QY 3377 ACAGCGACGACGAG 3390
DB 122 ACAGCGACGACGAG 135

RESULT 14
ADS31200
ID ADS31200 standard; DNA; 653 BP.
XX
AC ADS31200;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genome high complexity repeat found in the HIRA gene #233.
XX
KW Human; ds;
KW histone cell cycle regulation defective; S. cerevisiae homologue A; HIRA;
KW high complexity repeat; in situ hybridisation; Southern blot; disorder;
KW chromosome breakpoint; inherited genetic disease; neoplastic disorder;
KW chromosome 22; DiGeorge syndrome; Velo-Cardio-facial syndrome.
OS
XX Homo sapiens.
XX
FN US2003224356-A1.
XX
PD 04-DEC-2003.
XX
PF 14-MAY-2001; 2001US-00854867.
XX
PR 16-MAY-2000; 2000US-00573080.
XX
PA (KNOOL/) KNOOL J H M.
PA (ROGA/) ROGA P K.
XX
PI KNOOL JHM, Rogan PK;
XX
DR WPI; 2002-062378/08.
XX
PT Single copy genomic hybridization probes for detecting specific nucleic
PT acid sequences in sample by in situ hybridization useful for detection of
PT acquired or inherited genetic diseases.
XX
PS Example 1; SEQ ID NO 233; 30pp; English.
XX
CC The invention relates to a nucleic acid hybridisation probe comprising a
CC labelled, single copy nucleic acid of at least 50 nucleotides, which
CC will hybridise to a deduced single copy sequence interval in target
CC nucleic acid (TNA) of known sequence. The single copy sequence is deduced
CC by comparing the target nucleic acid (e.g. a disease causing gene) with a
CC collection of high and low complexity repeat sequences as found in the
CC genome of the organism from containing the target nucleic acid. The probe
CC is generated by PCR on the target sequence. The probe is essentially free

CC of blocking nucleic acid sequences which will hybridise to repeat
CC sequences within the genome of which the TNA is a part, and is labelled
CC with a label selected from fluorochrome-responsive labels, fluorochromes,
CC calorimetric chemical, conjugated proteins, antibodies, antigens and
CC their mixtures. The probe is useful in a hybridisation method, where the
CC hybridisation method is from in situ hybridisation, Southern blot, and
CC other methods in which nucleic acid is immobilised, where the method
CC further comprises selecting a single copy nucleic acid which will
CC hybridise to a duplication or triplication sequence domain. The probe is useful
CC for determining the existence of previously unknown repeat sequence
CC families in a genome. The method comprises reacting a labelled probe with
CC the genome, causing the probe to hybridise and ascertaining if the probe
CC hybridises to the genome at more than three preferably ten different
CC locations as a determination of new repeat sequence family, where the
CC determining step comprises selecting the single copy sequence from a
CC duplication or triplication sequence domain. The probe is useful for
CC determining a chromosome breakpoint and is useful in the fields for
CC cytogenetics and molecular genetics for determining the presence of
CC specific nucleic acid sequences in a sample of eukaryotic origin, e.g.
CC the probes may be used to analyse specific chromosomal locations by in
CC situ hybridisation as a detection of genetic or neoplastic disorders.
CC diseases especially for detection of genetic or neoplastic disorders.
CC Unlike prior art techniques, the probe permits more precise chromosomal
CC breakpoint determinations by in situ hybridisation. The genomic sequence
CC comprising the human HIRA gene (histone cell cycle regulation defective,
CC S. cerevisiae, homologue A) was analysed for single copy sequence
CC intervals for use as probes of the invention. HIRA is located on
CC chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and
CC Velo-Cardio-facial syndromes. The present sequence is a high complexity
CC repeat found within the human genome used to analyse the HIRA gene for
CC repeat regions. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?docid=20030224356.

SO Sequence 653 BP; 200 A; 139 C; 132 G; 176 T; 0 U; 6 Other;

Query Match 3.7%; Score 126.2; DB 7; Length 653;
Best Local Similarity 70.9%; Pred. No. 4.7e-15;
Matches 222; Conservative 1; Mismatches 81; Indels 9; Gaps 4;

QY 119 AGCTGACCTCTATTTATTTACAGTGTGTTATTTACCAATCAGCGCAACA-TGAATTA 177
DB 2 AGTGTATCTCTATTTATTTACAGTGTGTTATTTACCAATCAGCGCAACA-TGAATTA 61

QY 178 GTGATTTCTGAACCACTGCCCTAGGGAGACTACAGGTTAGATTCCTGCAAGCTCTGG 237
DB 62 GCGAATCTGAACCACTGCCCTAGGGAGACTACAGGTTAGATTCCTGCAAGCTCTGG 121

QY 238 TC--ACGTTTTTGTAAACAATCAATTAATTAACCTGTTTGTGTGCAATTTCTGTTTA 294
DB 122 TCACAACAATTTCAATCAACGATCAATTAATTAACCTGTTTGTGTGTTTCA 181

QY 295 AATATCTTTATTTATTAACGTAATCTGTAATTTCTCAATTTGTTCAACCAAGGC 354
DB 182 AGACNNC-TTATTTATTTATTAATTTGTAATTTATTAACCTGTAACCAAGGC-- 238

QY 355 CCTATTAATGGAAGCCCTGAATGAAGTTACATTAACACACATTTGTTTATGAGAA 414
DB 239 --CACTATTAATCAAGCTGTAATGAAGTTATTAACACACATTTTCTCATTAAGTA 296

QY 415 AATATTTTCTTC 427
DB 297 CATCAAGCTTC 309

RESULT 15
AA206449
ID AA206449 standard; DNA; 2742 BP.
XX
AC AA206449;
XX
DT 08-DEC-1999 (first entry)
XX

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 31, 2005, 09:15:07 ; Search time 382 Seconds

(without alignments)
14520.885 Million cell updates/sec

Title: US-09-743-312D-1_COPY_7103_10492

Perfect score: 3390

Sequence: 1 Tgctgctccgaagccatag.....atgagacagcgagcagaag 3390

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 240568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1459.2	43.0	3435	US-09-329-685A-1	Sequence 1, Appli
2	1459.2	43.0	3435	US-09-901-511-1	Sequence 1, Appli
3	525	15.5	2482	US-09-626-959D-1	Sequence 1, Appli
4	494.8	14.6	1537	US-09-626-959D-2	Sequence 1, Appli
5	422.4	12.5	649	US-09-354-409-1	Sequence 1, Appli
6	333.8	9.8	1537	US-09-626-959D-2	Sequence 2, Appli
7	259.4	7.7	396	US-09-354-409-4	Sequence 4, Appli
8	174.4	5.1	200	US-09-626-959D-3	Sequence 3, Appli
9	150	4.4	158	US-09-354-409-6	Sequence 6, Appli
10	143.6	4.2	156	US-09-354-409-8	Sequence 7, Appli
11	141.8	4.0	1196	US-08-891-789B-1	Sequence 8, Appli
12	134.4	4.0	1196	US-08-891-789B-1	Sequence 1, Appli
13	132.8	3.9	1240	US-08-891-789B-3	Sequence 1, Appli
14	126.2	3.7	653	US-09-573-080A-233	Sequence 233, App
15	126.2	3.7	2743	US-08-525-596B-13	Sequence 13, Appli
16	126.2	3.7	2743	US-09-177-860A-13	Sequence 13, Appli
17	126.2	3.7	2743	US-09-378-238-13	Sequence 13, Appli
18	126.2	3.7	2743	US-09-451-501-13	Sequence 13, Appli
19	126.2	3.7	2743	US-09-629-938-13	Sequence 13, Appli
20	126.2	3.7	2743	US-09-686-344-13	Sequence 13, Appli
21	126.2	3.7	2743	US-09-686-344-13	Sequence 1, Appli
22	126.2	3.7	2743	US-09-485-046-3	Sequence 1, Appli
23	119.2	3.5	35609	US-09-949-016-17370	Sequence 5, Appli
24	118.4	3.5	799	US-09-354-409-5	Sequence 12414, A
25	114.8	3.4	94830	US-09-949-016-12414	Sequence 12414, A
26	114.8	3.4	94847	US-09-949-016-16336	Sequence 16336, A
27	104.6	3.1	2676	US-08-525-596B-11	Sequence 11, Appli

28	104.6	3.1	2676	US-09-177-860A-11	Sequence 11, Appli
29	104.6	3.1	2676	US-09-378-238-11	Sequence 11, Appli
30	104.6	3.1	2676	US-09-451-501-11	Sequence 11, Appli
31	104.6	3.1	2676	US-09-629-938-11	Sequence 11, Appli
32	104.6	3.1	2676	US-09-686-344-11	Sequence 11, Appli
33	104.6	3.1	2676	US-09-686-344-11	Sequence 5, Appli
34	104.6	3.1	2676	US-09-485-046-1	Sequence 1, Appli
35	103	3.0	2676	US-08-891-789B-5	Sequence 1, Appli
36	100.6	3.0	601	US-09-949-016-163913	Sequence 15773, A
37	98.8	2.9	76264	US-09-949-016-15773	Sequence 12245, A
38	97.6	2.9	113100	US-09-949-016-12245	Sequence 15078, A
39	94.4	2.8	114426	US-09-949-016-15078	Sequence 76, Appli
40	93.8	2.8	932	US-09-573-080A-76	Sequence 11968, A
41	93.6	2.8	174259	US-09-949-016-11968	Sequence 14259, A
42	93.6	2.8	174262	US-09-949-016-14259	Sequence 32, Appli
43	93.4	2.8	66524	US-09-949-001-32	Sequence 38, Appli
44	93.4	2.8	66525	US-09-949-001-38	Sequence 14356, A
45	93.2	2.7	41617	US-09-949-016-14356	

ALIGNMENTS

RESULT 1	US-09-329-685A-1	Application US/09329685A
Sequence 1, Appli	Patent No. 6284882	
GENERAL INFORMATION:		
APPLICANT:	Abbott Laboratories	
APPLICANT:	Wu-Wong, Jinsun R.	
APPLICANT:	Wang, Jiahong	
TITLE OF INVENTION:	THE MYOSTATIN GENE PROMOTER AND	
TITLE OF INVENTION:	INHIBITION OF ACTIVATION THEREOF	
FILE REFERENCE:	6542 US 01	
CURRENT APPLICATION NUMBER:	US/09/329, 685A	
CURRENT FILING DATE:	1999-06-10	
NUMBER OF SEQ ID NOS:	3	
SOFTWARE:	FastSeq for Windows Version 4.0	
SEQ ID NO 1		
LENGTH:	3435	
TYPE:	DNA	
ORGANISM:	Homo sapiens	
US-09-329-685A-1		
Query Match	43.0%; Score 1459.2; DB 3; Length 3435;	
Best Local Similarity	72.7%; Pred. No. 0;	
Matches 2472; Conservative	0; Mismatches 743; Indels 187; Gaps 38;	
QY	1 Tgctgctccgaagccataggaaggttgcaaatcccgacatgagagaaacacgcttc	60
DB	119 TTCTATTTCTGGGGCTCAGAGGTTGCAATCAGTGCACCTGGAGAGAAATTAACCTTA	178
QY	61 TTGAATTTCTTGAATCCTTTAATTCATCAATGCTGACCTCCGA-----GATTGG	114
DB	179 CTTAAATTTCTTCAATTCCTTCAACCATCAATCAATGCTTCTTGAAGAGTTGATTAG	238
QY	115 ATAAGAGCATCTCATTTATTCAGAGGTATTTTCAACCATCAATCAATCAATCAATCA	173
DB	239 ATATAGTCATATTCCTTCAATTCATGATGATGATGATGATGATGATGATGATGATGAT	298
QY	174 ATAAGTATTCCTTCAATTCATTCAGAGGTATTTTCAACCATCAATCAATCAATCAATCA	233
DB	299 GTTAGCAAAATTTGAACTATTTGTTCCAGAGGAAACAGAGGTTAGTTGACCTCTG---	355
QY	234 CTGGTACGTTTGTGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	293
DB	356 GTCTAATCAATTTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	415
QY	294 AAAATCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	353
DB	416 CAAATA-CATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	470
QY	354 CCTATTATTAATGAGAGCCCTGATGAGAGCTTACATTAACACATTTGTTTCTATGAGGA	413

Db 471 CAGACATTAACCTGAGGCTGAATGATGCTTATCTAGACA---TGATTTTTCATGAG 527
Qy 414 AAAATTTTCTCTGATGCTGCGACAGGCTTCTGTTAAATTTGGACAAATATACA 473
Db 528 AACTTTTTCCTTATGAGCATATCAAGCTT-----TTAAATTTGTGTAAATATACA 582
Qy 474 TAAATGAAATGACTATTTATTTAACTGATTTTAACTGATGATGAGTCAATTAAGTA 533
Db 583 ACATTTAGCA-----TCTTAACCATTTTAACTTTAGTTGATGCTTGGATTTAGCA 633
Qy 534 CACTTACCTTGTGCTGATGATGATGATCAACATTCATCTAGAACTTTTATCTTCTG 593
Db 634 CATTCACATGTTGTATACATACACACATTTATCTTGAGAAATTTTATCTTCCCA 693
Qy 594 AATTGAAAATCTGCTATTTATGACAGTAGCTTTCCCATCACACCGACCTTCTC 653
Db 694 GACTG-AAACTCTGATCTATCACACAGTAACCTACCTCA-----GTGCTCAC 743
Qy 654 CAGCCCCCGGACACACATGCTTCTTCATGACAGTATCTGTGCTGCTAGAAACA--- 710
Db 744 CAGTCTCTGGACACACATGCTATCTTTCATGCTGCTTCTGTGATTTAGAAACATAG 803
Qy 711 -----CAGCCCTTCAACATCACCTGCTTGGGGGCACTGT---AAGCAACAGATC 756
Db 804 GAACACACGACGACGTCAGACATATGCTTGGGGGCACTTTTAAAGCAAAATCAAT 863
Qy 757 ACTCCCTACCGGACCAAAATGACACAAATATATAAAGCATGTGGCATATGCAATGC 816
Db 864 AAGAGACCAATATMAAAAGAGACAAATATGTGAAACATGTGCACTAAATACCGG 923
Qy 817 AAAAAGGGTGTGCT--AAGTATAGGGCTGAAC---AAGGACAGAAATGATAGT 871
Db 924 AAAAAAGGAATTTTGTATATGATATAGATGATGATGATGATGATGATGATGATGAT 983
Qy 872 TGACCTGACCTGGATCTGTGTGTGG--AAGCTCAATTTTCAATTTTCTGTGAT 929
Db 984 TGACCTGATCTGGGAACCTGTGCTGATGATGATGATGATGATGATGATGATGAT 1043
Qy 930 ACGCAAAATGCTTATMAAGCATGTGATGATGATGATGATGATGATGATGATGAT 989
Db 1044 GCGACGATGACTGATMAAA---TGCTAGGAATTTGATTTCAATGATMAATTTTATG 1099
Qy 990 CAAGCATTAATGTCAGACAGGATGATGATGATGATGATGATGATGATGATGAT 1047
Db 1100 AAGTGGCAATTTTGTATATGATGATGATGATGATGATGATGATGATGATGAT 1159
Qy 1048 TGATATCTTTTGTCTTTATTTCCAGGATGATGATGATGATGATGATGATGATGAT 1093
Db 1160 TGACAACTTTTGTCTTTTCCAGGATGATGATGATGATGATGATGATGATGATGAT 1219
Qy 1094 CACAGGCTGCTTACACATTTTCTAGAGCTAGGCTATGCTAGTGG----- 1144
Db 1220 AACCTGGCTGCTTGAATAATTTTCTGCTCAACAGGCTTGAAGAGTGGCTTGTGAC 1279
Qy 1145 ---TAGCTGGCAATTACAACTGAAGCAGTTCTAGTTCTAGTTCTAGTTCTAGTT 1201
Db 1280 TGATATTTGGACGCTATGACCTGAGCACTTCTAGTTCTAGTTCTAGTTCTAGTT 1339
Qy 1202 TAAATCTCAACCCCTCTGATGATAACAGAGCTAAGTACTCAAGTACAGTATCAGTCA 1261
Db 1340 TAAATCTCAACCCCTCTGATGATAACAGAGCAAGCACTCAAAATGCCAGTTATCAATTA 1399
Qy 1282 CTATCTATATGACAGGACCTGATCTACAAATTTTATGATATTTATGATGATGATGAT 1321
Db 1400 CTATCTATATGACAGGCTGATCTACAAATTTTATGATATTTATGATGATGATGAT 1459
Qy 1332 CAACCTGATGAGAGGCTGAAGTTAGAG---AAGTATCTCATTTATATACACA 1376
Db 1460 CAACCTGATGAGAGGCTGAAGTTAGAGAGTTAGATATCTCATTTATATACACA 1519
Qy 1377 GT-----GGCAAACTGAGATCTGAACTCAGGTCTATCACTCAGAGACTGAGATC 1428

Db 1520 GTTAGAAGTGCAAAAGTTGAGATTTGAATCTAGCTATATCTGACCTCAGAGCTGAGTTC 1579
Qy 1429 CAAATGCTACACAAATTTCTATCAAGTTAAAGGAAAAAGATTTGATTTGCTCAGAG 1488
Db 1580 TCAATTTCACTGCTATCAATTTCTAGACATATTAAGGAAAAAGTTGATCTTACTGAGAC 1639
Qy 1489 TGTATAGGGCATATGTTACA---ATTATACATTTACAAAGATTTATATGTTGAAAAATA 1545
Db 1640 TGTATAGTCAATGTTTAAACATGATCATTAATTTGAAAGATTTTACATGAAATATG 1699
Qy 1546 AATTATCAAAACAAATTAATTTTAAAGCTGATCTTAATACGCTCCGCAACAAAGACT 1605
Db 1700 AATTATCAAAACAAATTAATTTTAAAGCTGATCTTAATATGTTCCATTAACAAAGACT 1759
Qy 1606 ATCTGAATCTTCAAGGACATCTGTTGTGTGTGTTGTTTCTTAATCTTAAATGATGAG 1665
Db 1760 ATCTGAATCTTCAAGGACATCTGTTGTGTGTGTTGTTTCTTAATCTTAAATGATGAG 1818
Qy 1666 CAATCTAATGATATGTAAGGCA---TTTTTCTCAAGAGATGATGATGATGATGATGAT 1724
Db 1819 CAATCTAATGATATGTAAGGCAATTTTCTCAAGAGATGATGATGATGATGATGATGAT 1878
Qy 1725 AATTGATGAAAATGATTAATTTTCAAGCTGATGATGATGATGATGATGATGATGATGAT 1784
Db 1879 AATTGATGAAAATGATTAATTTTCAAGCTGATGATGATGATGATGATGATGATGATGAT 1938
Qy 1785 GTAATTTAGTATCAATGATGAGAAA---GTATGATCTTAAATTAATATTTATGATGAT 1841
Db 1939 GTAATTTAGTATCAATGATGAGAAAATGATGATGATGATGATGATGATGATGATGAT 1998
Qy 1842 ACACATGATCTCTTGG--GGAATAAATGATGATGATGATGATGATGATGATGATGAT 1900
Db 1999 GCACTGATCTCTTGGAAAAAAGATGATGATGATGATGATGATGATGATGATGATGAT 2058
Qy 1901 TCCCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1960
Db 2059 TCCCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2118
Qy 1961 CTGCTATATTTGTTAGTTTATATGATGATGATGATGATGATGATGATGATGATGAT 2020
Db 2119 CTGCTATATTTGTTAGTTTATATGATGATGATGATGATGATGATGATGATGATGAT 2178
Qy 2021 CTAAAGTTGAGATTTGATCTCTTCAAGGCTCACTCAATTTAGAAACAAAGCCCA 2080
Db 2179 CTAAAGTTGAGATTTGATCTCTTCAAGGCTCACTCAATTTAGAAACAAAGCCCTA 2238
Qy 2081 TTTCTGCTCCCAACCTTACCCCAATCCCTGCAAGGTCTGCTCCGCTGCAATGAT 2140
Db 2239 TTTCTGCTCCCAACCTTACCCCAATCCCTGCAAGGTCTGCTCCGCTGCAATGAT 2298
Qy 2141 GAAACTGCAAAAGAACTAGAGGTGCGACAGTACTAGAAAGTAAAGTGAAGTGA 2200
Db 2299 AGAAATTTGGCAAAAGGCTGCAACA-----TA 2325
Qy 2201 GCACATACGTAAGAAAGCAAAATGAGGACCTTCAATGATGATGATGATGATGATGATGAT 2260
Db 2326 TCGCAGATTTGGGAAACAAAGGTCACCCCTTATCAAGATGCTCTTCTCTTTAT 2385
Qy 2281 GTTTCACAAAGCTCCGATATATTTTACAGAGGATGATGATGATGATGATGATGATGAT 2312
Db 2386 GTTTCACAAAGCTCCGATATATTTTATGAAATGATGATGATGATGATGATGATGATGAT 2445
Qy 2313 TCTTTTACACGTAAGGCTGAGAAACCTTGTATACCATCATTAATTTACATATCTTC 2372
Db 2446 AATTTTACGCTGAGAAATCTAGGCAATCTGATATCTGCGATGCGATGATTAATA 2505
Qy 2373 TAAATCTATATGTTATTTTGAAG---TCAATATGCTGACATATCTCTTTGTA 2426
Db 2506 TTTATTTATCTATATGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2565
Qy 2427 TAAACATGAAAAACATATCTCTGACATATTAATCTGCACTTTAGATGATGATGAT 2486
Db 2566 TAAACATG-AAAAACATCTTCTGAGTTATGTTATCTGATCTTTAGATGATGATGAT 2622

```

Qy 2487 AACTTAATACAGTCAATGAACTGAATACAAATTTTCATATGAAATAAAGATAT---- 2542
Db 2623 ---ATTAATAGACCTCAGTCAAAAGTTCAGTAAATTTTCATATTAATTAAGACATGAAA 2679
Qy 2543 -TATTAAAGTAATTCATGAGCAATTTAATTA-AAGTGAATTTTCATATGCTT 2600
Db 2680 CTATGTAAATAATTAATTCATGACATATGTTATATTAACATGATCTCAATATTTTACT 2739
Qy 2601 AAGAATTATTCAGGGAACAAGTTCTCAATTAATTAACAGAAATCTTTTACTGTATTC 2660
Db 2740 AAGAAATTTAGTCA-GAATAACAAGTTCTCAATTAATTAAGTGAATTTCTCAATAGTATC 2798
Qy 2661 ACAGCTTTTCAATTTAGTCT-TTCCGTAATTAATCTGATTTTCTAATTAATCAAGACTA 2719
Db 2799 ATAAATCTTAATTTTAATTCAGGCTCTCCATATTTTATTTTCCATTTACTGGCACAATA 2858
Qy 2720 AAATAATTTAATTAACAATAAATTAATTTT-TACTTCAATGCTTACTTAATAGTA 2778
Db 2859 AAAATAATTTAATTAACAATAAATAATTTTCTACTTCAAAATCTTGCCATAACAATA 2918
Qy 2779 TAAATCATTTTATTTCTGAGGGAACAATATCACTTTTATAGTATGAAGTAAAT 2838
Db 2919 TAAATCATTTTATTTCTGAGGGAAGTAATTTTCT-ATAATTTAATTAATGTAATTAAT 2977
Qy 2839 TAAATTTATTTCACTTAATTTAATTAATTTTAAAGTTTCAATATAAAGATGATAGATC 2898
Db 2978 TAAATTTCACTTAATTTAATTAATTAATTAAGTGTGTGAGATTAAGTAAATTAAGTATC 3037
Qy 2899 TAAATTTATTTATTTGTTAATTAATTAATTTTAAATTTTCCATGATCAATACAGCTTT 2958
Db 3038 ---AGTTAATTAATTTATTTCCAAATGATGATGATTTTCCAAATGATGATGATTTT 3093
Qy 2959 ATTAATTCATAGATTTATCTTTTAAAGATGATGATGATGATGATGATGATGATGATGATG 3018
Db 3094 ATTAATTTATTTATTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3153
Qy 3019 TAACAAAATTTATTTTGTGTCAGTGTGTCAGACAGAGGTTTAACTTGTGACAGAGATTT 3078
Db 3154 TAACAAAATTTATTTGTCGTGTCAGTGTGTCAGACAGAGGTTTAACTTGTGACAGAGATTT 3213
Qy 3079 CATTTGTGAGCAAGACCAATCAAGATCCGACAGACATTTGTCTCATCAAAATTTGAAT 3138
Db 3214 CATTTGTGAGCAAGACCAATCAAGATCCGACAGACATTTGTCTCATCAAAATTTGAAT 3273
Qy 3139 ATTAAGGCACTTGAATATGATTAAGATTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 3198
Db 3274 ATTAAGGCACTTGAATATGATTAAGATTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAG 3333
Qy 3199 ACTGGGCAAGCATTTACGTTGGCTTGGCTTACTCAAAAGCAAAAGAAAGTAAGAAAGA 3258
Db 3334 ACTGTACATGATTAATTTTCTGTCGATTTACTCAAAAGCAAAAGAAAGTAAGAAAGA 3393
Qy 3359 AGAAGTAAGAACAGGAAAGATTTGATTTTAATTAACCT 3300
Db 3394 AGAAGTAAGAACAGGAAAGATTTGATTTTAATTAATC 3435

```

RESULT 2

```

; Sequence 1, Application US/09901511
; Patent No. 639312
; GENERAL INFORMATION:
; APPLICANT: Abbot Laboratories
; APPLICANT: Wu-Hong, Jinhyun R.
; APPLICANT: Wang, Jiahong
; TITLE OF INVENTION: THE MYOSTATIN GENE PROMOTER AND
; FILE REFERENCE: 6542, US, DI
; CURRENT APPLICATION NUMBER: US/09/901,511
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 09/329,685
; PRIOR FILING DATE: 1999-06-10

```

```

; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-901-511-1

```

```

Query Match 43.0%; Score 1459.2; DB 3; Length 3435;
Best Local Similarity 72.7%; Pred. No. 0;
Matches 2472; Conservative 0; Mismatches 743; Indels 187; Gaps 38;

```

```

Qy 1 TGTGCTCCGAAGCCATGGAAGGTTGCAATCCCGGACATGAGAGAGAAACAAGTTC 60
Db 119 TTCTATTTCTGGGGCTCGAGAGGTTGCAATCAGTGCACATGAGAGAGAAATTAATCTTA 178
Qy 61 TTGAATTTCTTGAGTACCTTTAATTTCAATTTCTGACCTCCGA-----GATTGG 114
Db 179 CTTAATTTCTTGAGTACCTTTAATTTCAATTTCTGACCTCCGAAGGTTGATTAG 238
Qy 115 ATAGAGTCACTCATTTATTCAGAGTGTATGTTTACCAATCACTGCAACA-TGA 173
Db 239 ATATGATCAATTTCTCCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 298
Qy 174 ATAGATTTCTGGAACCACTGCCCTAGAGGAACTACAGATTAGATTTCCGTCAGCT 233
Db 299 GTTGAACAATTTATGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 355
Qy 234 CTGGTCACTTTTGTGTTAACCATCAATTAATTAATTTTGTGTTGTTGTTGTTGTTT 293
Db 356 GTCAATCAATTTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 415
Qy 294 AAAATTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 353
Db 416 CAATTA-CATTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 470
Qy 354 CCTATTATTAAGGAGCCGTAATGAACTTACATTAACAACATTTTCTATGAGA 413
Db 471 CAGACATTAATCAAGGCTGATATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 527
Qy 414 AAAATTTTCTTCAATTTCTGTCAGAGCTTTCTTCTTAAATTTGTGACAAATTAACA 473
Db 528 AACTTTTCTTCAATTTCTGTCAGAGCTTTCTTCTTAAATTTGTGTAATTAATTAACA 582
Qy 474 TAAATGAACTGATTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 533
Db 583 ACATTTAGCA-----TCTTAACATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 633
Qy 534 CACTTACCTTGTGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 593
Db 634 CATTTACATTTGTGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 693
Qy 594 AATTGAAATCTGCAATTTATTTGACAGTATTTTCCCATCAACCGCACCTTCTC 653
Db 694 GACTG-AAACTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 743
Qy 654 CAGCCCGCCGACCAATCTTCTTCTTCAATGACATGATCTGTCGTGAGAAACA--- 710
Db 744 CAGTCCCGGACCAATCTTCTTCTTCAATGACATGATCTTCTGTCGTGAGAAACAATG 803
Qy 711 -----CAGCCTTCAACATTTAGCTTGGGGGCACTGT-----AAGCAACGATC 756
Db 804 GAACACGACGACAGACGTCAGACATTTGCTTTGGGGGCACTTTTAAAGCAAAATCAAT 863
Qy 757 ACTCCCTACGCGCACCAATGACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 816
Db 864 AAGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 923
Qy 817 AAAAGGCTGCTGCT--AAGTATGAGGCTGAAG--AAGCAGAGATTTGATAGT 871
Db 924 AAAAGGAAATTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 983
Qy 872 TGACCTGAGCTGGATCTGTGTGTGG--AAGCTCAAAATTTTCAATTTGTTGTGTGAT 929

```

Db 984 TGACCTCATCTGGGAACTGTGGCGTACAGATAGAGACTCAAAATTTTGGCCGCTGTGTAAT 1043
 Qy 930 AGCGACAATCTCTTAATAAAGACCTTAAGATGATTAAGAATTAATCTCAG 989
 Db 1044 GCCACGCGAATACGAAAAA---TGCTAGGAATATTAAGTTTCATAATTAATTTTATAGC 1099
 Qy 990 CAAGACATAAATGTGCAAGACGGGATCCATGAATTAAGACACTGACCAT--GTGAAA 1047
 Db 1100 AAGTAGGCAATTTGTGTAATACAGAAATCATGAATTAAGAATTTGTCTGTAGCGAAA 1159
 Qy 1048 TGATATCTTTTGTTCCTTAATTCAGGACAGTAAGAGAAAGCG-----T 1093
 Db 1160 TGACAACTTTGTTCATTTTTCAGAGGTGAATGAATTAAGAGCTGGAACATGCTTGT 1219
 Qy 1094 CACAGGCTGCTTACACCATTTTACTAGAGAGCTAGCTATGTCAAGCG----- 1144
 Db 1220 AACCTGGCTGCTTGAATAATTTTACTTGCTCAAGGCTAGAAAAGTAGGCTTTGTGAC 1279
 Qy 1145 ---TAGCTGGCAATTAACAAGTGAAGCAGTTCTAGTTATGAGAGATGAATTTAACCA 1201
 Db 1280 TGATATTTGGAGCTATAGCTGAAGCAGTTCTAGTTATGAGAGATGAATTTTAAACA 1339
 Qy 1202 TAACTCAACCCCTCTGCAATGAACAGAGACTAAGTACTAAGTACAGATTATCAGTCA 1261
 Db 1340 TAACTCAACCCCTCTGCAATGAACAGAGCAGCAGTCAATGAGCAGTTATCAATTA 1399
 Qy 1262 CTATCTATATAGACAGGCACTGTACTCAACATTTACATGATTTATGAATTAATCATGCC 1321
 Db 1400 CTATCTATATAGACAGGCTGCAATTTACAGCAATTTATCATGCAATTAATTAATATCCCC 1459
 Qy 1322 CAACACTCTATAGAGAGCTGAAGTTAGAG---AAGTATCTATTCATTTATTAACA 1376
 Db 1460 CAACACTCTATAGAGAGCTGAAGTTAGAGAGTTAGAGTTAGTATCTATTCATTTATTAACA 1519
 Qy 1377 GT-----GGCAACTGAGATCTGAATCTCAGGTCTATCCAACTCCAGAGCTGAGATC 1428
 Db 1520 GTTAGAAGTGGCAAGTTGAGATTTGAACCTCAGTCTATCTGACTCCAGAGCTGAGTTC 1579
 Qy 1429 CCAATTCGTACCAATTTCTAATCAAGTTAAAGGAAAAAGATTGATTTGCTCAGAG 1488
 Db 1580 TCAATTTCAACGTGCTATACAAATTTCTAAGCATTTAAAAAAGTTGACTTACCTGGAAC 1639
 Qy 1489 TGTATAGGGGCAATAGTTACA---ATTATTAATTATCAAAAGATTTATATGTTGAAAAATA 1545
 Db 1640 TGTATAGATGATGCTTATCAATGATCATACATTTGAAAGATTTACATTTGAAAAATG 1699
 Qy 1546 AATTTATCAACAAATAAACTTTATTAAGCTGATCTAATCTGCTCCGCAACAAAGCT 1605
 Db 1700 AATTTATCAACAAATAAACTTTGAAAGCAGATCTAATATTTGCCATTAACAAAGAT 1759
 Qy 1606 ATCTGAATCTTTCAGGAGATCTGTTGTCTGCTGTTTCTTAAATCTTTAATGATGG 1665
 Db 1760 ATCTGAATCTTTCAGGAGATCTGTTGTCTGCTGTTTCTTAAATCTTTAATGATGG 1818
 Qy 1666 CAATCTATAGCATTAATGTAAGGCA--TTTTTCTCAAGAGATGATACCTTTAAGA 1724
 Db 1819 CAATCTATAGCATTAATGTAAGGCACTTTTCTCAAGAGATGTAATCTCTTTAAGA 1878
 Qy 1725 ATTTGATGAATGCAATTAATCTTTCAGGCTACTGATCTGATTTTATGTCAGTGAAGCA 1784
 Db 1879 ATTTGATGAATGCAATTAATCTTTCAGGCTACTGATCTGATTTTATGTCAGGAGCA 1938
 Qy 1785 GTAAATTAATGTAATGTAAGGCA---GTAGTGAATCAAAATTAATTTGATATGA 1841
 Db 1939 GTAAATTAATGTAATGTAAGGCAATGTAAGTGAATTTTAAATTAATTTGATATGA 1998
 Qy 1842 ACCACGTGCACTCTCTGG--GGAATAAGTATGATTAATCTCTTAGAGATCTTAGCT 1900
 Db 1999 GGCACGTGATTTCTTTGGAAAAAAGTAATGATTAATCTCTTAGAGATCTTAGCT 2058
 Qy 1901 TCCCAAAAGTATGAAAAATTAATCTCTGTGGCTGGAACAGCTTCTGTCTTGG 1960

Db 2059 TCCCAAAAGGATGAGAAAAAGAAATCTCTTGGCTTGAATAATCTTCTGTCTTG 2118
 Qy 1961 CTGCGATTAATTTGTTTATGATTTTAAATGATTAATGATTAAGCTTGTGGCTCCCAAG 2020
 Db 2119 CTGCGATTTGTTGTTGATGCTCTTTAATGATTAATGATTAATGATTTGTGGCTCCCAAG 2178
 Qy 2021 CTAAAGTTGAGAGTTTATGATTTCCCTTACAGAGGCACTTCAATTTTGAAGAACAAAGCCCA 2080
 Db 2179 CTAAAGTTGAGAGTTTATGATTTCCCTTACAGAGGCACTTCAATTTTGAAGAACAAAGCCCA 2238
 Qy 2081 TTTCTGCTCCCAAGACCTTACCCCAATCCCTGCAAGGTGCTGCTCCGCTCAAAATGA 2140
 Db 2239 TTTCTGCTCCCAAGACCTTACCCCAATCCCTGCAAGGTGCTGCTCCGCTCAAAATGA 2298
 Qy 2141 GAACTGGCAAGGAGTACTAGAGAGGTGCGACAGTACTAGAGATGAATAATGACTA 2200
 Db 2299 AGAATTTGGCAAGGGTGGCAACA-----TA 2325
 Qy 2201 GCACACTAAGTGAAGAGCAAAAAATGCGACCTTCAATGATGTGTTCTTCCCTTCT 2260
 Db 2226 TCGCAGATTTGGAAACAACAAAGTCAACCCCTTATCATGATGCTCTTCTTAT 2385
 Qy 2261 GTGTTCAATGCTCCGATTAATTTACAGAGGTAGATA-----ACTACATTTT 2312
 Db 2386 GTGCTCAATTAATTTCTGATTAATTTTATGAGAAATGATCTGCACTTTTATCTCTG 2445
 Qy 2313 TCTTTTACACTGGAAGGCTGAGAAAACTTGTATCCATCATTAATTCATATCTTC 2372
 Db 2446 ATATTTACTGCGAAATCTGAGCAATCTGTATATCTGCGATCCAGTTATATAA 2505
 Qy 2373 TAACTCATTTATGTTATTTCTAGA-----TCAATAGCTGACAAATCTCTTTGTA 2426
 Db 2506 TTCAATTAATTAATGTTATGTTCAAGATTTTCTTACAGCTGGGATTAACCTCTGTGTA 2565
 Qy 2427 TAAACATGAAGAAAAACATCTCTGAGCAATTAATCTGCACTTTAGATGGAAGT 2486
 Db 2566 TAAACATG--AAAAACATCTCTGAGTATGTTAATCTGCACTTTAGATGGAAG-- 2622
 Qy 2487 AACTTAATTAATGCTGAATGGAATCAATTTTCAATGATTAATGAAT-- 2542
 Db 2623 ---ATTAATGACCTGACAAAGTTCAGTAAATTTTCAATTAATTAAGACATGAA 2679
 Qy 2543 -TATTTAAAGTAAATCCATGAGCAATTAATTA--AAGTAGATTTTCAATGCT 2600
 Db 2680 CTATGTAATAATTAATCCATGACAAATGTTAATTAATTAATTAATTAATTAATTAAT 2739
 Qy 2601 AAGAAATTAATCAAGGAAACAGTTTCTCAATTAATGACAGAAATCTTTACTAGTATC 2660
 Db 2740 AAGAAATTAATCA--GAAACAAAGTTTCTCAATTAATGATTAATTAATTAATTAATTAAT 2798
 Qy 2661 ACAGCTTTTCAATTAATGTC--TTCGTAATAATCTGATTTTCTAATTAATTAATTAAT 2719
 Db 2799 ATTAATCTTAATTTTAATTAAGCTTCTTCAATTTTATTTTCTTAATTAATGACATA 2858
 Qy 2720 AAAATTAATTAATTAATTAATTAATTTT--TACTTCAATGCTTAATTAATTAAT 2778
 Db 2859 AAAATTAATTAATTAATTAATTAATTTTCTTCAATTAATTAATTAATTAATTAAT 2918
 Qy 2779 TAAATCAATTTTATTTTCTGAGGAAAGCATATCACTTTTAAATGTAAGTAAAT 2838
 Db 2919 TAAATCAATTTTATGTTTGGAGAAATATTTT--AATTTTAAATTAATTAATTAAT 2977
 Qy 2839 TAAATTAATTAATTAATTAATTTTAAAGTTTCAATTAATTAATTAATTAATTAAT 2898
 Db 2978 TAAATTAATTAATTAATTAATTAATTAATTAATTTTCCATGATTAATTAATTAAT 3037
 Qy 2899 TAAATTAATTAATTAATTAATTAATTAATTTTATTTTCCATGATTAATTAATTAAT 2958
 Db 3038 ---AGTTAATTAATTAATTAATTAATTAATTTTCTTCAATTAATTAATTAATTAAT 3093
 Qy 2959 ATTAATTAATTAATTAATTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 3018
 Db 3094 ATTAATTAATTAATTAATTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 3153

QY 3019 TAACAAATTAAGTGGTGAAGTCTTGACAGACAGGGTTTAACTCTGACAGCAGATT 3078
 DB 3154 TAAACAATATAGCTGCTGGTGAAGTCTTGACAGACAGGGTTTAACTCTGACAGCAGATT 3213
 QY 3079 CATTTGGAGCAGACGCCAATCAAGATCCGACGACACTTGTCTCATCAAGTTGGAAAT 3138
 DB 3214 CATTTGGAGCAGACGCCAATCAAGATCCGACGACACTTGTCTCATCAAGTTGGAAAT 3273
 QY 3139 ATAAAGCCCACTTGGAAATACGATTAAGATGACAGTGTGTGGCAAGTTGTCTCTAG 3198
 DB 3274 ATAAAGCCCACTTGGAAATACGATTAAGATGACAGTGTGTGGCAAGTTGTCTCTAG 3333
 QY 3199 ACTGGCAGCAGATTACGCTTGGCTTGGCGTCTACTCAAAAGCAAAAGAAAGTAAAGGA 3258
 DB 3334 ACTGACATGATTAATTAATTTTGGCTTGGCATTACTCAAAAGCAAAAGAAAGTAAAGGA 3393
 QY 3259 AGAAGTAAGAACAGAGGAAAGATTTGATTTTAAACC 3300
 DB 3394 AGAAGTAAGAACAGAGGAAAGATTTGATTTTAAATC 3435

RESULT 3

US-09-626-959D-1
 ; Sequence 1, Application US/09626959D
 ; Patent No. 6617440
 ; GENERAL INFORMATION:
 ; APPLICANT: Findlay, Robert
 ; TITLE OF INVENTION: MYOSTATIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AND
 ; TITLE OF INVENTION: METHODS FOR ITS USE
 ; FILE REFERENCE: PCI0448A
 ; CURRENT APPLICATION NUMBER: US/09/626,959D
 ; CURRENT FILING DATE: 2000-07-27
 ; PRIOR APPLICATION NUMBER: 60/146,540
 ; PRIOR FILING DATE: 1999-07-30
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2482
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-09-626-959D-1

Query Match 15.5%; Score 525; DB 4; Length 2482;
 Best Local Similarity 61.5%; Pred. No. 1,2e-112;
 Matches 1256; Conservative 0; Mismatches 665; Indels 122; Gaps 21;

QY 1188 GATGAATTTAAACATATCTCAACCCCTCTGCATGAAACAGAGACTAAGTACTCAAGTA 1247
 DB 521 GAACCACTTGAGATTAATTTGAAAGTAAGGCTTGAACAAAGAGCAAGCCCTTCTGCT 580
 QY 1248 CCAGTTACAGTCACTTACTATATG-----AAGGCACTGTACTCAACATTTAAT 1299
 DB 581 TCAAGTATTAATTAATCTATGAAAGGAGACTACATTTAGCTATATATGCTAAATTAAT 640
 QY 1300 GTATTATTAATTAATCATGCCCCCACTATGAGAAAGTGAAGTTAGAGATATC 1359
 DB 641 GCCTCAAAACCCCTTATGTTGAGAAATCTAAGATTAAGAGAGTATGATGCGCTCTT 700
 QY 1360 TCATTCATTAATTAACAGATGCAAACTGAGATCTGAATCTCAAGTCTATCACTCCAGGA 1419
 DB 701 TGTCACTCACTTGAAGAGGCAAAATGAGATTTGAATCTCAAGTCTTATTTGACTCTTCA 760
 QY 1420 CCTGAGATCCCAATTTGCTACACAAATCTAATCAATTTAAAGGAAAGAAAGTTGATTT 1479
 DB 761 TCTGAGCTCAACATGCGAGTCAAGTCTAATAAAGAAAGAAAGTCAAGATCA---ATTT 816
 QY 1480 GCTCAGAGTATATAGGCGCATATGTAC---AATTATTAATTAATCAAGATTTATATGT 1536
 DB 817 CCTGAGATTAATTAAGCAGCATGTGTATGATTAATTAATTAATTAATTAATTAATTAAT 876
 QY 1537 TGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1596

DB 877 AGAATAATTAATTTCCAGACAAATCTGATTTATAGCCTGCTCTAATATTTCTTTGA 936
 QY 1597 ACAAAGACTATCTGAATTCCTTCAGGAGCATGCTGGTTGTGTGCTGTTTCTTAATCTT 1656
 DB 937 --TAAAGGGCGAGATCACTCAGGGTGTGCTGCTTGTGTGCTGTTTCTTCAATCTT 994
 QY 1657 AATGATGGCAATCTAATGATTAATGTAAGGCCA-----TTTTTCTCAAGATATAG 1711
 DB 995 AATGATGGCAATCTAATGATTAATGTAAGGCCACTTTTCTTCTCAAGATATAG 1054
 QY 1712 ATACCTCTTAAGATTTGATGAAATGCTATTAACCTTTCAAGCTTACTGATTTGATTTA 1771
 DB 1055 ATGCTCTTAATAAATTTGATGAAATGCTATTAACCTTTCAAGCTTACTGATTTGATTTA 1114
 QY 1772 GTGCACTGAGCAGTAATTT---AGTGTCAATGTGGCAAGTGTGACTTAATAATTA 1828
 DB 1115 GTTCACTGAGCAGTAATTTGGGTGTATACGTATCAGAGAAATGTGTGACTTAATAATTA 1174
 QY 1829 ATATTGATTAATGAACCACTGCACTCTCTTGGGG-----AAAAAGTAATGATTAAC 1880
 DB 1175 ATATTGATTAATGAACCACTGCACTCTCTTGGGGGTGTGTAAAGGAGATTAATTAATTA 1234
 QY 1881 TCTCTTAAGAGTCTTAACTTCCCAAAAGTATGAAATTAATCTCTGCTGCTG 1940
 DB 1235 ATCTTAAGAGTCTTCAAGCTTCCCAAAAGGAGAGGAA-----CTCTGCTGCTG 1285
 QY 1941 GAAACAGCTTCTGTTCTGCTGCTGCTATTTGTTAGTTTAAATGATTAATTTGATTT 2000
 DB 1286 GAAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343
 QY 2001 AGACCTTGTGCTCTCCAAAGCTTAAGTGAAGTTGATTTGATTTGATTTGATTTGATTT 2060
 DB 1344 CGATTTGTGCTCTCTTAAGGCAAGGAGTGAAGTTGATTTGATTTGATTTGATTTGATTT 1403
 QY 2061 TTGAAGAAACAAAGCCCAATTTCTGCTCCCAAGCTTACCCCAATCTCTGCTGCTG 2120
 DB 1404 TCAGAGAAACAAAGCCCAATTTCTGCTCCCAAGCTTACCCCAATCTCTGCTGCTGCTG 1462
 QY 2121 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2180
 DB 1463 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1518
 QY 2181 GAAAGTAAGAAATGCACTGACACTGAGAGAGCAAGAAATGCGACCTTCAATGA 2240
 DB 1519 -----GGAAAGCCCAACAAAGACACCCCTCTTACA 1547
 QY 2241 TGTGTTCTCTTCCCTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 2300
 DB 1548 ---TGGCACTTGTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603
 QY 2301 ACTACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2351
 DB 1604 ---TTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1660
 QY 2352 ATCAATAAATTAATCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2411
 DB 1661 TACTTCAAAATTAATCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1720
 QY 2412 ATATCTCTTGTGTAATAAACAATGAAACAATCTCTGAGCAATTAATCTGCAAC 2471
 DB 1721 GGAATTCCTGTTGCAATCAAAAGAAAGAAAGAAAGCAACCTGCTTCTTGTGTGT 1780
 QY 2472 TTTAGATAGGAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2531
 DB 1781 ATCTCTGTAATGAAGAAATGAAGCAATTAATTAATTAATTAATTAATTAATTAATTA 1840
 QY 2532 ATAAAGATTAATTT---TTAAAGTATTCATGAGCAATTAATTAATTAATTAATTAATTA 2586
 DB 1841 GTCAAGCACTTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1900
 QY 2587 TTTCAATTAATGTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2646
 DB 1901 TTGTAATTAATGTTTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1960

QY 2647 CTTTACTAGTACAGAGCTTTTCATTAGTCTCTCTGATAAATCTGATTTTCTAA 2706
 DB 1961 TTTTACAGAGGTGACAACTCAGCTTCTTTAAATTAAGTACGCTTCTTCAAGTTTAC 2020
 QY 2707 TTATACAGACTAAATAATTTATATATACAAATAAATTTATTTTACTTCAATGCTT 2766
 DB 2021 TTCT-----CTAATTACCGACACTTAACGCAATTTTTCCTCCAAATATTA 2068
 QY 2767 ACTTAAATAGTATAAATCATTTTATTTTCTGAGGAAACCATATCACTTTTAACTA 2826
 DB 2069 GTTTTATGACAAAACAGCAGCTCCAGCTCAAGATTAACATTTTCTA-TTTTAAACA 2127
 QY 2827 TGAAGTAAATTAAGATTTTATCACTAAATTAATTTTAAAGTTTCAATATAAG 2886
 DB 2128 CAAATCTAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2180
 QY 2887 ATGATTAAGATCTAAGTATATGTTATGTTAAATTAAGTTTAAATTTTTCGACATCA 2946
 DB 2181 ATTATTAAGCTTTAAGTACAGTTATATTAATGTT-ACAGAGCTTCAATTTTCAATGTC 2239
 QY 2947 CATACAGCTTTATTTATTCATAGATTTTTCCTTTAAGAAAGTACTCAATGAATCAGCT 3006
 DB 2240 CATATATCTTTCATGATTTGGGAGATTTATTTTCAATTTATGAAGTCAATGAATCAGCT 2239
 QY 3007 CACCTTGAAGTAAACAAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3066
 DB 2200 TGCCCTGACGTGTAACAAAGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2359
 QY 3067 GACAGCGAGTTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3126
 DB 2360 GACAGCGAGTTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2419
 QY 3127 CAAAGTTGAATTAATAAAGCCACTTGTGATTAACATTAATTAATTAATTAATTAATTA 3186
 DB 2420 CTAAGTTGAATTAATAAAGCCACTTGTGATTAACATTAATTAATTAATTAATTAATTA 2479
 QY 3187 GTT 3189
 DB 2480 GTT 2482

RESULT 4

US-09-626-959D-2
 ; Sequence 2, Application US/09626959D
 ; Patent No. 6617440
 ; GENERAL INFORMATION:
 ; APPLICANT: Findlay, Robert
 ; TITLE OF INVENTION: MYOSTATIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AN
 ; FILE REFERENCE: PCI0448A
 ; CURRENT APPLICATION NUMBER: US/09/626,959D
 ; CURRENT FILING DATE: 2000-07-27
 ; PRIOR APPLICATION NUMBER: 60/146,540
 ; PRIOR FILING DATE: 1999-07-30
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 1537
 ; TYPE: DNA
 ; ORGANISM: Sus scrofa
 US-09-626-959D-2

Query Match 14.6%; Score 494.8; DB 4; Length 1537;

Best Local Similarity 83.7%; Pred. No. 1.2e-105;

Matches 676; Conservative 0; Mismatches 102; Indels 30; Gaps 9;

QY 2395 AGATCAATAGCTGCAATATCTCTTTGTAATTAACATGAAACAAACATCTCTGAG 2454
 DB 747 AATTAATGAATGACATTAATCTCTTTGTAATTAACATGAAACAAACATCTCTGAG 806
 QY 2455 CAATATTAATCTGCACTTAGATGAAATTAATTAATTAATTAATTAATTAATTAATTA 2514

DB 807 CAATATTAATCTGCACTTAGATGAAACAAACAAACATCTCTGAG 855
 QY 2515 ATCAATTTTCAATGATTAATAAAGATTAATTTTAAAGTATTCATGACAAATTAATA 2574
 DB 856 GCAGATTTTCACTGATTAATAAAGATTAATTTTAAAGTATTCATGATTAATAA 915
 QY 2575 TTA-AAGTAAATTTTCAATGATTAATAAAGATTAATTTTAAAGTATTCATGATTAATAA 2633
 DB 916 GAATTAATGATTAATTAATGATTAATAAAGATTAATTTTAAAGTATTCATGATTAATAA 975
 QY 2634 TATAGCGAAATTTTCAATGATTAATAAAGATTAATTTTAAAGTATTCATGATTAATAA 2687
 DB 976 TATAGCGAAATTTTCAATGATTAATAAAGATTAATTTTAAAGTATTCATGATTAATAA 1035
 QY 2688 ATAAATCTGATTTT-TCATTAATTAACAAGTAAATAATTTTAAATTAATAAATAAAT 2746
 DB 1036 TTAATCTGATTTTCTGATTAACAAGTAAATAATTTTAAATTAATAAATAAAT 1095
 QY 2747 TATTTTACTGAAATCTTAATTAATAAAGTAAATAATTTTAAAGTATTCATGATTAATAA 2806
 DB 1096 TCTTTTCTCAATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1155
 QY 2807 --GATATCACTTTTAAAGTAAAGTAAATAATTTTAAAGTATTCATGATTAATAA 2864
 DB 1156 AGACATTTCACTTTTAAAGTAAAGTAAATAATTTTAAAGTATTCATGATTAATAA 1211
 QY 2865 TTTTAAAGTTTCAATTAATAAAGTAAATAATTTTAAAGTATTCATGATTAATAA 2921
 DB 1212 TTTTAAAGTTTCAATTAATAAAGTAAATAATTTTAAAGTATTCATGATTAATAA 1271
 QY 2922 AAAGTTTAAATTTTTCGATGTCATCAATCAAGCTTTTATTAATTAATTAATTTTCTTT 2981
 DB 1272 ATAGATTTTAAATTTTTCGATGTCATCAATCAAGCTTTTATTAATTAATTAATTTTCTTT 1331
 QY 2982 TAAAGATGTCATTAATAAAGTAAATAATTTTAAAGTATTCATGATTAATAA 3041
 DB 1332 TATGAATGTCATTAATAAAGTAAATAATTTTAAAGTATTCATGATTAATAA 1391
 QY 3042 TGTGACAGACAGGTTTAACTCTGACAGGAGATTCATTTGTGAGGACCAATCA 3101
 DB 1392 TGTGACAGACAGGTTTAACTCTGACAGGAGATTCATTTGTGAGGACCAATCA 1451
 QY 3102 CAGATCCGACGACCTTGTCTCATCAAAAGTTGCAATTAATAAAGCCACTTGAATACAG 3161
 DB 1452 TAGATCCGACGACCTTGTCTCATCAAAAGTTGCAATTAATAAAGCCACTTGAATACAG 1509
 QY 3162 TATTAAGATTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3189
 DB 1510 TATTAAGATTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1537

RESULT 5

US-09-354-409-1
 ; Sequence 1, Application US/09354409
 ; Patent No. 6555672
 ; GENERAL INFORMATION:
 ; APPLICANT: Li-fang Liang
 ; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR PROMOTER AND USES
 ; FILE REFERENCE: MTN-027
 ; CURRENT APPLICATION NUMBER: US/09/354,409
 ; CURRENT FILING DATE: 1999-07-15
 ; EARLIER APPLICATION NUMBER: 60/092,865
 ; EARLIER FILING DATE: 1998-07-15
 ; EARLIER APPLICATION NUMBER: 60/123,270
 ; EARLIER FILING DATE: 1999-03-08
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 649
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-354-409-1

Query Match 12.5%; Score 422.4; DB 4; Length 649;
Best Local Similarity 82.0%; Pred. No. 6,6e-89;
Matches 536; Conservative 0; Mismatches 111; Indels 7; Gaps 4;

```
QY 2652 ACTGATACAGCGCTTTTCATTAAAGTC-TTCCTGAATTAATCGTATTTCTAATAT 2710
DB 1 ACTGATACATATATCTTAACCTTTAATTCAGGCTCTCTAATTTTATTTTCTTAATAC 60
QY 2711 ACAAGACTAAATAATTAATATAACAATAAATTAATTTT-TCCTCAATGCTTACT 2769
DB 61 TTGGCACTTAATAATATTTAATACAAATATAAATATTTTCTAATTAATCTTGGC 120
QY 2770 TAAATAGTAATAAATCAATTTTATTTTCTGAGGAAAGCATATCACTTTTAACTATCA 2829
DB 121 TAAACAAATAATAAATCAATTTTATTTTGAAGAAATATTTTC-ATATTTAAATATGT 179
QY 2830 AGCTGAATTAAGATTTATCTTAATATTAATTTTAAAGTTTCACATATAAAGATG 2889
DB 180 AGTATAAATTAATAATGACTTAATTAATTAACAATAAGATGTGTGAGGATTAAGTA 239
QY 2890 AATAAGATCTAAGTATATGTTATTTGTTAATAAGTTTAAATTTTTCGATGTCACAT 2949
DB 240 TTTAAG----TACGTTTATATATTTGCGACATGACTTTGTTTTCAAATGTCAAA 295
QY 2950 ACAAGCTTTATTTATCATAGATTTATTCCTTTTAAAGATGCTCAATGAATCACTCAC 3009
DB 296 ATATCTTTTATTTATTTAGATTTATTTCTTTATGAAGTACTCAATGAATCACTCAC 355
QY 3010 CCTGACCTGTAACAATAATCTGTTTGCTGACCTTGACAGACAGGCTTTAACTCTGAC 3069
DB 356 CTTGACCTGTAACAATAATCTGTTTGCTGACCTTGACAGACAGGCTTTAACTCTGAC 415
QY 3070 AGCGATTCATTTGAGGACAAAGCAATCACAGATCCGACGACCTGTCTCATCA 3129
DB 416 AGCGATTCATTTGAGGACAAAGCAATCACAGATCCGACGACCTGTCTCATCA 475
QY 3130 AGTTGAATTAATAAAGCACTTGGAATACAGTATTAAGATTTCACTGTGTGCAAGT 3189
DB 476 AGTTGAATTAATAAAGCACTTGGAATACAGTATTAAGATTTCACTGTGTGCAAGT 535
QY 3190 GTCTCTGACCTGGGACGACATTAAGTTGGCTTGGCGTTACTCAAAAAGCAAAA 3249
DB 536 GTCTCTGACCTGGGACGACATTAAGTTGGCTTGGCGTTACTCAAAAAGCAAAA 595
QY 3250 GTAAAAAGAAAGTAAGAACAAGGAAAGATGTGATTTGATTTAAACCATG 3303
DB 596 GTAAAAAGAAAGTAAGAACAAGGAAAGATTTGATTTGATTTAAACCATG 649
```

RESULT 6
US-09-626-959D-2/c
; Sequence 2, Application US/09626959D
; Patent No. 6617440
; GENERAL INFORMATION:
; APPLICANT: Findlay, Robert
; TITLE OF INVENTION: MYOSTATIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AN
; TITLE OF INVENTION: METHODS FOR ITS USE
; FILE REFERENCE: PCI0448A
; CURRENT APPLICATION NUMBER: US/09/626,959D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/146,540
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Sus scrofa
US-09-626-959D-2

Query Match 9.8%; Score 332.8; DB 4; Length 1537;
Best Local Similarity 63.9%; Pred. No. 7.8e-68;

Matches 555; Conservative 0; Mismatches 302; Indels 11; Gaps 3;

```
QY 2485 GTAACCTAATCTAGTCATTTGAACGAATATACATTTCTATATGAATTAAGAATATTA 2544
DB 1223 GAACCTTTAAATTTGTAATTTAAATTAAGTATTTCTTTACACTTATCTTAAAGTTG 1164
QY 2545 TTTAAAGTAATTCAGTACGACATTTAAATTAATTAAGTATTTCTATATGTTGAAGA 2604
DB 1163 AAATGCTTTTCCCAAAAAAATAAATGAATTTTACATTTATTTGACAAACATTTGAGG 1104
QY 2605 ATTATTTGAGGAAACAAGTTTCTCAATTTATAGAGAAAATCTTTACTATGATACAG 2664
DB 1103 AAAAAAGATTTTATTTCTGTTTAAATTTATTTTAACTGCTGTGTAATTCAGAAATAC 1044
QY 2665 TCTTTCAATTAAGCTTCTGTAATTAATCTGATTTTCTAATTTATACAACTAAAAAT 2724
DB 1043 AGATTTAAATTTAGGAAGCTGGAATTA----ATTTAAGATTTGATACATAGTAAATA 988
QY 2725 AATTTAATATAACAATAAATTAATTTTACTTCAATGCTTACTTAATATGATATAAT 2784
DB 987 TATTCAGCTAATATTTGAGAAACCTGTTTCTGACTAAATTTCTAGTACATTAATGAANA 928
QY 2785 CATTTATTTTCTGAGGAAAGCATATCACTTTTAAATGATGAAGTAAATTAAGAT 2844
DB 927 TCATACCTTAATCTGTTATTAATTAACAATGAATTAATTTTAAATTAATCTTTTATTCAG 868
QY 2845 TTATTCATTAATTAATTAATTTTAAAGTTTCAATATTAAGATTAAGATTAAGTCAAGT 2904
DB 867 TGAATATGCTGCTCACTTTTCACTGATTTTCTCTATCTTAAGTTGCAAGTTAATTAAT 808
QY 2905 TATATGTTATTTGTAATAAGTTTAAATTTTTCGATGTCATATACAGCTTTAATTAAT 2964
DB 807 GCTCAGAAAGATGTTTCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 748
QY 2965 CATGATTTATTTCTTTTAAAGTATGCAATATGATAGCTCACTTGTACTGTAACA 3024
DB 747 TGTGAATTTATTTCTTTTAAAGTATGCAATATGATAGCTCACTTGTACTGTAACA 688
QY 3025 AATACCTGTTGGTGACTTGTGACAGACAGGCTTTTAACTCTGACAGGAGATTCATTTGT 3084
DB 687 AATACCTGTTGGTGACTTGTGACAGACAGGCTTTTAACTCTGACAGGAGATTCATTTGT 628
QY 3085 GGAGCAAGGCAATCAAGATCCGACGACACTTGTCTATCAAAAGTTGGAATTAATA 3144
DB 627 GGAGCAAGGCAATCAAGATCCGACGACACTTGTCTATCAAAAGTTGGAATTAATA 570
QY 3145 AGCACTTGAATACAGTATTAAGATTTCACTGTGTGCGCAAGTTGTCTTCAAGCTGG 3204
DB 569 AGCACTTGAATACAGTATTAAGATTTCACTGTGTGCGCAAGTTGTCTTCAAGCTGG 510
QY 3205 CAGGATTAACGTTTGGCTTGGCTTACTCAAAAGCAAAAAGTAAGTAAGGAAGAT 3264
DB 509 CAGGATTAACGTTTGGCTTGGCTTACTCAAAAGC-----AAAAGTAAGGAAGGAAT 455
QY 3265 AAGAAACAAGGAAAGATTTGATTTGATTTTAAACATGCAAAAACCTGCAATCTCTGT 3324
DB 454 AAGAAACAAGGAAAGATTTGATTTGATTTTAAACATGCAAAAACCTGCAATCTATGTT 395
QY 3325 TATATTTACTTATTTTATGCTGATTTGTTG 3352
DB 394 TATATTTACTTATTTTATGCTGCTATCTG 367
```

RESULT 7
US-09-354-409-4
; Sequence 4, Application US/09354409
; Patent No. 6555672
; GENERAL INFORMATION:
; APPLICANT: Li-fang Liang
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR PROMOTER AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MTN-027
; CURRENT APPLICATION NUMBER: US/09/354,409

;; CURRENT FILING DATE: 1999-07-15
;; EARLIER APPLICATION NUMBER: 60/092,865
;; EARLIER FILING DATE: 1998-07-15
;; EARLIER APPLICATION NUMBER: 60/123,270
;; EARLIER FILING DATE: 1999-03-08
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 4
;; LENGTH: 396
;; TYPE: DNA
;; ORGANISM: Mus musculus
US-09-354-409-4

Query Match 7.7%; Score 259.4; DB 4; Length 396;
Best Local Similarity 82.3%; Pred. No. 6.2e-51;
Matches 312; Conservative 0; Mismatches 61; Indels 6; Gaps 1;

QY 2920 ATAAAGTTTAAATTTTCGATGTCACATACAGCCTTATTTATTCATAGATTTATTCCT 2979
DB 18 ACACAGACTTCATTTATTCATATGTCATATATCTTCATGATTTGGGATTTATTTCA 77
QY 2980 TTTAAGATGTCATATGATCAGCTCACCTTGTACTGTATACAAATACCTTTGGTGA 3039
DB 78 TTTATGAAGTAGTCAATGAATGACCTTGCCCTGACCTGTAAACAAATACCTTGGTGA 137
QY 3040 CTGTGACAGACAGGGTTTAACTCTGACAGCGAGATTTATGAGCAAGCCAT 3099
DB 138 CTGTGACAGACAGGGTTTAACTCTGACAGCGAGATTTATGAGCAAGCCAT 197
QY 3100 CACAGATCCGACGACACTTGTCTCATGAAAGTTGAAATATTAAGCCACTTGGATAC 3159
DB 198 CATGATCCTGACGACACTTGTCTCATGAAATATTAAGCCACTTGGATAC 257
QY 3160 AGTATTAAGATTCACCTGTGTGGCAAGTTGTCTCTCAGACTGGGCAAGCATTAAGTT 3219
DB 258 AGTATACAGGACCTCCCTGGCGGAGGTGTCTCTCGACGCTCATGCACTAATATTT 317
QY 3220 GCGTTGGGCTTACTCAAAAGCAAAAGAAAGTAAAGAGAAAGTAAAGCAAGGAAAA 3279
DB 318 CACTTGGCTTACTCAAAAGCAAAAGAAAGTAAAGAGAAAGTAAAGCAAGGAAAA-----AAAA 371
QY 3280 GATTGTATTTGATTTTAAAA 3298
DB 372 GATTGTGCTGATTTTAAAA 390

RESULT 8
US-09-626-959D-3
;; Sequence 3, Application US/09626959D
;; Patent No. 6617440
;; GENERAL INFORMATION:
;; APPLICANT: Findlay, Robert
;; TITLE OF INVENTION: MOSTANTIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AN
;; FILE REFERENCE: PC10448A
;; CURRENT APPLICATION NUMBER: US/09/626,959D
;; CURRENT FILING DATE: 2000-07-27
;; PRIOR FILING DATE: 1999-07-30
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 3
;; LENGTH: 200
;; TYPE: DNA
;; ORGANISM: Mus musculus
US-09-626-959D-3

Query Match 5.1%; Score 174.4; DB 4; Length 200;
Best Local Similarity 92.0%; Pred. No. 3.3e-31;
Matches 184; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 2990 AGTCAATGAATCAGCTCACCTTGACTGTAAACAAATACCTGTTGGTGAATTTGACAG 3049
|||||

DB 1 AGTCAATGAATCAGCTTGCCCTTGACTGTAAACAAATGCTGCTGTGATCTGTGACAG 60
QY 3050 ACAGGGTTTAACTTGAAGCAGAGATTCATTTGTGAGCAAGCCATCAGATCCC 3109
DB 61 ACAGGGTTTAACTTGAAGCAGAGATTCATTTGTGAGCAAGCCATCAGATCCC 120
QY 3110 GACGACACTGTCTCATCAAAAGTTGAAATATTAAGCCACTTGGATACATTAAG 3169
DB 121 GACGACACTGTCTCTCATCAAAAGTTGAAATATTAAGCCACTTGGATACATTAAG 180
QY 3170 ATTCACTGTGTGGCAGTT 3189
DB 181 ACTCCTGGCGGTGGCAGTT 200

RESULT 9
US-09-354-409-6
;; Sequence 6, Application US/09354409
;; Patent No. 655672
;; GENERAL INFORMATION:
;; APPLICANT: Li-fang Liang
;; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR PROMOTER AND USES
;; FILE REFERENCE: MTN-027
;; CURRENT APPLICATION NUMBER: US/09/354,409
;; CURRENT FILING DATE: 1999-07-15
;; EARLIER APPLICATION NUMBER: 60/092,865
;; EARLIER FILING DATE: 1998-07-15
;; EARLIER APPLICATION NUMBER: 60/123,270
;; EARLIER FILING DATE: 1999-03-08
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 6
;; LENGTH: 158
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-354-409-6

Query Match 4.4%; Score 150; DB 4; Length 158;
Best Local Similarity 96.8%; Pred. No. 1.5e-25;
Matches 153; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2985 GAAGTAGTCAATGAATCAGCTCACCTTGACTGTAAACAAATACCTGTTGGTGAATTT 3044
DB 1 GAAGTAGTCAATGAATCAGCTCACCTTGACTGTAAACAAATACCTGTTGGTGAATTT 60
QY 3045 GACAGACAGGGTTTAACTTGAAGCAGAGATTCATTTGTGAGCAAGCCATCAGATCCC 3104
DB 61 GACAGACAGGGTTTAACTTGAAGCAGAGATTCATTTGTGAGCAAGCCATCAGATCCC 120
QY 3105 ATCCGACGACACTGTCTCATCAAAAGTTGAAATATTA 3142
DB 121 ATCCGACGACACTGTCTCATCAAAAGTTGAAATATTA 158

RESULT 10
US-09-354-409-7
;; Sequence 7, Application US/09354409
;; Patent No. 655672
;; GENERAL INFORMATION:
;; APPLICANT: Li-fang Liang
;; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR PROMOTER AND USES
;; FILE REFERENCE: MTN-027
;; CURRENT APPLICATION NUMBER: US/09/354,409
;; CURRENT FILING DATE: 1999-07-15
;; EARLIER APPLICATION NUMBER: 60/092,865
;; EARLIER FILING DATE: 1998-07-15
;; EARLIER APPLICATION NUMBER: 60/123,270
;; EARLIER FILING DATE: 1999-03-08
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 7

TELEFAX: (416) 863-2653
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-891-789B-3

Query Match 3.9%; Score 132.8; DB 3; Length 1240;
Best Local Similarity 98.5%; Pred. No. 3.6e-21;
Matches 134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3255 AGAAGAAGTAAGAACAGGAAAGATGATGATTTAAACCATGCAAAAACGCA 3314
DB 1 AGGAAAAAGTAAGAACAGGAAAGATGATGATTTAAACCATGCAAAAACGCA 60
QY 3315 AATCTGTTTATATTATCTATTTATGCTGATTTGCTGCGCCAGTGAATGA 3374
DB 61 AATCTGTTTATATTATCTATTTATGCTGATTTGCTGCGCCAGTGAATGA 120
QY 3375 GAACAGCGAGCAGAG 3390
DB 121 GAACAGCGAGCAGAG 136

RESULT 14
US-09-573-080A-233
; Sequence 233; Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOUL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 233
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(653)
; OTHER INFORMATION: mer82
; NAME/KEY: misc feature
; OTHER INFORMATION: n is a, c, g or t
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Mlodoszewska, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-233

Query Match 3.7%; Score 126.2; DB 4; Length 653;
Best Local Similarity 70.9%; Pred. No. 9.7e-20;
Matches 222; Conservative 1; Mismatches 81; Indels 9; Gaps 4;

QY 119 AGCTACTCATATTATTCACAGGTATGTTCTACCATCATCGCAACA-TGAATA 177
DB 2 AGTATCTCTCATTAATCAACAGAGTATGTTCTATAAAGTACCGCGAGCACTGAATA 61
QY 178 GGTATTCCTGAACCACTGCCCTAGGGGAATTAAGAGTTAGATTCCTCAGCCTCTGG 237

DB 62 GCGAATACGAGAACCATCGCTCTAGAGAAATACAGATTAGGTTCCCTGACGCTCTGG 121
QY 238 TC---ACGTTTGTATACCATCAATTAATACCTGTTGTGTCATTTCTGTTTA 234
DB 122 TCAACATTTTCAATCAACCGATCAATATTAACCTTTTATGTTGTGTTTAA 181
QY 295 AATATCTTTATTAATACGTAAGTCTGTAATTTCTTCAACATTTGTTTCAACAAAGGC 354
DB 182 AGACNNC-TTATTTAATATATATGTTGATTTATTAACACTGAACACAGCAAG-- 238
QY 355 CTAATTAAGTGAAGCCCTGAATGAAGCTTACATACACACATGTTTCTATGAGAA 414
DB 239 -CACTTAACCTACGCTGCTGAATGAAGCTTATCTAACACACATATTTCTCATTAAGTA 236
QY 415 AATTTTCTCTC 427
DB 297 CATCAGCCCTTC 309

RESULT 15
US-08-525-596B-13
; Sequence 13; Application US/08525596B
; Patent No. 5827733
; GENERAL INFORMATION:
; APPLICANT: Huynh, Thanh
; APPLICANT: Lee, Se-Jin
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,596B
; FILING DATE: 19-SEP-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07762
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., Ph.D, John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07265/075001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2743 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: Human GDF-8
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 59...1183
US-08-525-596B-13

Query Match 3.7%; Score 126.2; DB 1; Length 2743;
Best Local Similarity 91.2%; Pred. No. 1.7e-19;
Matches 134; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 3243 AAGAAAGTAAGAAAGAAAGTAAGAACAGGAAAGATTTGATTTTAAACCAT 3302

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 12:46:17 ; Search time 1218 Seconds
(without alignments)
16845.606 Million cell updates/sec

Title: US-09-743-312D-1_COPY_7103_10492
Perfect score: 3390
Sequence: 1 tgcgtccgcagcagccatagc.....atgagacagcgcagcagaag 3390

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5607317 seqs, 3026245999 residues
Total number of hits satisfying chosen parameters: 11214634

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCITS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1459.2	43.0	3435	9	US-09-901-511-1
2	525	15.5	2482	17	US-10-610-473-1
3	508.6	15.0	5790	15	US-10-251-115-54
4	494.8	14.6	1537	17	US-10-610-473-2
5	332.8	9.8	1537	17	US-10-610-473-2
6	224.8	6.6	942	13	US-10-027-632-120231
7	224.8	6.6	942	17	US-10-027-632-120231
8	190	5.6	2823	17	US-10-366-345-27
9	174.4	5.1	200	17	US-10-610-473-3
10	134.4	4.0	1196	15	US-10-251-115-1
11	132.8	3.9	1240	15	US-10-251-115-3

12	126.2	3.7	653	10	US-09-854-867-233	Sequence 233, App
13	126.2	3.7	653	19	US-10-786-970A-233	Sequence 233, App
14	126.2	3.7	2743	9	US-09-859-211-13	Sequence 13, Appl
15	126.2	3.7	2743	9	US-09-841-730-1	Sequence 1, Appl
16	126.2	3.7	2743	10	US-09-872-856-13	Sequence 13, Appl
17	126.2	3.7	2743	14	US-10-278-803-13	Sequence 13, Appl
18	126.2	3.7	2743	15	US-10-335-483-13	Sequence 13, Appl
19	126.2	3.7	2743	17	US-10-463-973-13	Sequence 11, Appl
20	126.2	3.7	2743	17	US-10-456-852-3	Sequence 3, Appl
21	126.2	3.7	2743	18	US-10-662-438-1	Sequence 1, Appl
22	126.2	3.7	2743	19	US-10-665-374-1	Sequence 1, Appl
23	104.6	3.1	2676	9	US-09-859-894A-10	Sequence 10, Appl
24	104.6	3.1	2676	9	US-09-859-211-11	Sequence 11, Appl
25	104.6	3.1	2676	9	US-09-841-730-3	Sequence 11, Appl
26	104.6	3.1	2676	10	US-09-872-856-11	Sequence 11, Appl
27	104.6	3.1	2676	14	US-10-278-803-11	Sequence 11, Appl
28	104.6	3.1	2676	15	US-10-335-483-11	Sequence 11, Appl
29	104.6	3.1	2676	17	US-10-463-973-11	Sequence 11, Appl
30	104.6	3.1	2676	17	US-10-456-852-1	Sequence 1, Appl
31	104.4	3.1	330973	13	US-10-087-192-1498	Sequence 1498, Ap
32	103	3.0	2676	15	US-10-251-115-5	Sequence 5, Appl
33	102.8	3.0	55005	18	US-10-719-993-6968	Sequence 6968, Ap
34	102.8	3.0	318760	18	US-10-719-993-6765	Sequence 6765, Ap
35	100.4	3.0	98439	19	US-10-741-600-17724	Sequence 17724, A
36	97.4	2.9	1182	13	US-10-027-632-258706	Sequence 258706,
37	97.4	2.9	1182	17	US-10-027-632-270981	Sequence 270981,
38	96.6	2.8	708	13	US-10-027-632-270982	Sequence 270982,
39	96.6	2.8	708	13	US-10-027-632-270981	Sequence 270981,
40	96.6	2.8	708	17	US-10-027-632-270982	Sequence 270982,
41	96.6	2.8	708	17	US-10-027-632-174714	Sequence 174714,
42	94.8	2.8	620	13	US-10-027-632-174714	Sequence 174714,
43	94.8	2.8	620	17	US-10-027-632-174714	Sequence 174714,
44	94.8	2.8	853	13	US-10-027-632-157095	Sequence 157095,
45	94.8	2.8	853	13	US-10-027-632-157096	Sequence 157096,

ALIGNMENTS

RESULT 1
US-09-901-511-1
; Sequence 1, Application US/09901511
; Publication No. US20010049435A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Wu-Wong, Jinhyun R.
; TITLE OF INVENTION: THE MYOSTATIN GENE PROMOTER AND
; TITLE OF INVENTION: INHIBITION OF ACTIVATION THEREOF
; FILE REFERENCE: 6542 US D1
; CURRENT APPLICATION NUMBER: US/09/901, 511
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 09/329, 685
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-901-511-1

Query Match 43.0%; Score 1459.2; DB 9; Length 3435;
Best Local Similarity 72.7%; Pred. No. 1.7e-263;
Matches 2472; Conservative 0; Mismatches 743; Indels 187; Gaps 38;

QY 1 TGCCTCCGAGCATAGAGGTTCAATCCCGACATGAGAGAAACAGTTC 60
DB 119 TTCTATCTGGGCGCTCAGAGAGTTCAATCATGCACTGAGAGAAATAACTTA 178
QY 61 TTGAATTTCTTGAGTACTTATTAATTCATTCAGTGCATCTCCGA-----GATTGG 114
DB 179 CTTAAATTTCTTCAGTCTTCTTCCACCAATTCATTAATCTGAGAGAGTGAATTAG 238

115 ATAGAGCTGACTCTATTATTCACAGTGTATGTTCTACCCATTCACCTGCCAACA-TGA 173
174 ATAGATGATTCCTGACCACTGCCCTAGAGGGAATCAAGATTTAGATCCCTCAGCCT 233
174 ATAGATGATTCCTGACCACTGCCCTAGAGGGAATCAAGATTTAGATCCCTCAGCCT 233
299 GTTAGCAAAATTTGAATGTTTCCAGAGGGAAGGAGGTTAGTTGAGCCTCTG--- 355
234 CTGGTCAAGTTTGTGTTAACCAATCAATTAATTAACCTTTGTTGTGTCATTTCTGTTT 293
356 GTCAATACATTTACATCACCAATCAATATATTAACCTTTGTTGTGTTATGTTATGTTT 415
294 AAAATCTTATTTAATAGTACGCTAATCTTCAACATTTGTTGTTGTTGTTGTTGTTGTT 353
416 CAATTA-CATATTTAATATATATTTGTTTCAATTTCACTGAACTCAAGCCA----G 470
354 CCTATTAACTGGAAGCCTGAAATGAAGCTTACATTAACACATTTGTTTCTATGAGA 413
471 CAGCACTAATACAGGCTGAATGATGCTTATCTAGACA---TGATTTTTCATGAG 527
414 AAAATTTTCTTCACTGCTGACAGCCTTCTGTTAAATTTGTGACAAATATACA 473
528 AACTTTTCCCTTAGGATATCACAGCCTT-----TTAAATTTGTGTTAATATACACA 582
474 TAACATGAATGACTATTTAATTTAATCTGATTTAATCTGATGAGTTCACTGAGCTTAAGTA 533
583 ACATTTGACA-----TCTTAACATTTTATAGTTTCAAGTTGCTGAGCTTAAGTA 633
534 CACTTACCTTCTGCTGGAATGATGATCAATTCATCTTGAATCTTTGATCTTCTG 593
634 CATTCACCTGTTGTGTAACATGACACATGATCTTCAAGAAATTTTATCTTCCCA 693
594 AATTGAAAATCTGCTATTTGACAGTACGTTTCCCAATCACACCGCACCTTCTC 653
694 GACTG-AAACTCTGATCTATCACACAGTAACTACCTCA-----GTGCTCAC 743
654 CAGCCCCCGGACACCAATCCTTCTTCAATGACATCATCTGCTGCTGAGAAACA--- 710
744 CAGTCTCTGGACACACATGCTACTTTCATTTGAGCTTCTCTGTGATGAGAAATAG 803
711 -----CAGCCTTCAACATCACTGCTTGGGGGCACTGT---AAGCAACGATC 756
804 GAACACACGAGACGTCAGACATATGCTTGGGGCATTTTTAAAGCAAAATCAAT 863
757 ACTCCTTACCGGACCAATGACACAAATATTAAGCATGCTGCTATGCTATGCTC 816
864 AAGAGACCAATTAAGAAAGAAAGCAAAATATGTAAGAAACATGCACTTAATAGAGG 923
817 AAAAGGGTGTGCT--AAGTATAGGGCTGAAC---AAGGACAGAAATGACTAGT 871
924 AAAAGGGAAATTTGTTTATGATAGAGCTGAACAAAGGCGAAACCTGCTTGT 983
872 TGACCTCAGCTGGGATCTGTGTGTGG--AAGCTCAAAATTTTCATTTGTTGTCAT 929
984 TGACCTCATCTGGGAACCTGTGCTCAATAGATCAAAATTTTGGCGCTGTGAT 1043
930 ACGCAAAATGCTTATTAAGACATGTAAGATGATTTGAAGTTAAGATTAATCTCAG 989
1044 GCCACGATGACTAAAAA---TGCTAGAAATTTGATTTCAATGAATAATTTTATG 1099
990 CAAGCATTAATGTCAGACAGGATCAGATTAAGACAGACCTGACAT--GTGAAA 1047
1100 AAGTGGCAATTTTGTATTAAGAAATCATGAATTAATGAATGTCTGTATAGGGA 1159
1048 TGATATCTTTGTTCTTTATTTCAAGGAGTAAAGAAAGCG-----T 1093
1160 TGACAACTTTGTTCTTTTTCAGAGTATTAAGATTAAGAGCTGGAATGTTCTGT 1219
1094 CACAGGCTGCTTACACATTTTATAGAGCTGAGCTATGCTAGTGG----- 1144
1220 AACCTGCTGCTTGAATAATTTTACTTCTACAGGCTTGAAGAGTGGCTTTGTGAC 1279

1145 ---TAGTGGCAATTAACAACTGAAGCAGTTTCACTGATGAGAGTAACTTAACCA 1201
1280 TGATTAATTTGGCACTATGACCTGAAAGAGTTTCACTGATGAGAGTAACTTAAGCA 1339
1202 TAATCTCAACCCCTCTGCTATGAAACAGAGCTAAGTACCAAGTCCATATATCTGCA 1261
1340 TAACTCAACCCCTCTGCTATGAAACAGAGCTAAGTACCAAGTCCATATATCTGCA 1399
1262 CTTACTATATGACAGGCACTGTAATCAATTTACATGATTTAATTAATGATGCCCC 1321
1400 CTTACTATATGACAGGCTGTAATCAATTTACAGCAATTTACATGATTTAATATCCCC 1459
1322 CAACACTCTATGAGGAAGCTGAAGTTAG---AAGTATCTCATCTATTTACACA 1376
1460 CAAAACCTTATGAGGAAGCTGAAGTTAGGAAGTTAGATTCATCTCATTTATACATA 1519
1377 GT-----GGCAACTGATCTGAACTCAGCTCATCTCACTCAGAGCTGAGATC 1428
1520 GTTAGAAGTGGCAAGTTGAGATTTGAATCAGCTCATCTGACTCAGAGCTGAGATC 1579
1429 CCAATGCTACACAAATCTATCAAGTTAAAGGAAAGAAAGATTTGATTTGCTCAGAG 1488
1580 TCAATTTCACTGCTATCAATTTCAAGCATATTAAGAAAGTTGACCTTACTTGAAC 1639
1489 TGATAGGGGCAATGTTTACA---ATTATACATTAACAAGATTTATGTTGAAAATA 1545
1640 TGATATGATGATGATGTTTACATGATCATTAATTTGAAAGATTTACATTTGAAAATG 1699
1546 AATTATCAACAAATTAATTTATTAAGCTGATCTTAATACGCTCGGACAAAGACT 1605
1700 AATTATCAACAAATTAATTTATTAAGCTGATCTTAATACGCTCGGACAAAGACT 1759
1606 ATCTGAATCTTCAAGGCACTGCTGTTGTGCTGTTTCTTAACTTTAATGATGG 1665
1760 ATCTGAATCTTCAAGGCACTGCTGTTGTGCTGTTTCTTAACTTTAATGATGG 1818
1666 CAATCTAATGATATGATTAAGGCA---TTTTTTCTCAAGAGATGATGATCTTTAAGA 1724
1819 CAATCTAATGATATGATTAAGGCA---TTTTTTCTCAAGAGATGATGATCTTTAAGA 1878
1725 AATTGATGAAGAAATGATTAATCTTCAAGGCACTGATGATTTATGATGATGAGCA 1784
1879 AATTGATGAAGAAATGATTAATCTTCAAGGCACTGATGATTTATGATGATGAGCA 1938
1785 GTAATTAATGATTAATGATGAGCA---GTAATGATCTTAAGAAATTAATATTTATGA 1841
1939 GTAATTAATGATTAATGATGAGCA---GTAATGATCTTAAGAAATTAATATTTATGA 1998
1842 ACACACTGACTCTTGTG--GGAAGAAAGTAAATGATTAATCTCTTGAAGTCTTACT 1900
1999 GCACTGATCTCTTGGAAAGAAAGTAAATGATTAATCTCTTGAAGTCTTACT 2058
1901 TCCCAAAAGTAAAGTAAAGTAAATCTCTGAGGCTGGAACAGCTTCTGTTCTG 1960
2059 TCCCAAAAGTAAAGTAAAGTAAATCTCTGAGGCTGGAACAGCTTCTGTTCTG 2118
1961 CTGCTATATTTGTTAGTTTATTAATGATTAATGATTAATGATTAATGATTAATGAT 2020
2119 CTGCTATATTTGTTAGTTTATTAATGATTAATGATTAATGATTAATGATTAATGAT 2178
2021 CTAAAGTTGAGATTTGATCTTCAAGAGCTCACTTAATTTAGAAACAAAGCCCCCA 2080
2179 CTAAAGTTGAGATTTGATCTTCAAGAGCTCACTTAATTTAGAAACAAAGCCCCCA 2238
2081 TTCTCTGCTCCAGCTTAAACCAATCCCTGCAAGTGTCTGCTCGGCTCAATATGA 2140
2239 TTCTCTGCTCCAGCTTAAACCAATCCCTGCAAGTGTCTGCTCGGCTCAATATGA 2298
2141 GAATCTGCAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2200
2239 AGAATTTGGCAAGAGGGGTGCAACA-----TA 2325
2201 GCACTACTAGAAAGCAAGAAATGAGGACCTTCAATGATGATGATGATGATGATGAT 2260

Db 2326 TCGGATTTGAGAAACAACAAAGTCACCCCTTATCATGATCTCTCTTTAT 2385
 Qy 2261 GTGTTCAATGCTCCGATATATATTAACAGAGGGTAGATA-----ACATATTTT 2312
 Db 2386 GTGCTCAATATTTCTGATATATATTAAGAAATAGATCTGACCTTTTACTCTGG 2445
 Qy 2313 TCTTTTACACTGGAAGGCTGAGAAACCTTGTATCCATCATATAAATTCATCTTC 2372
 Db 2446 ATATTTACTGCTGAAATCTGAGCAACGTATATATCTCGCATCCAGTTATATAA 2505
 Qy 2373 TAAGCATTTATGTTATTTCTAAG------TCAATAGCTGACATATCTCTTTGA 2426
 Db 2506 TTCAATTACTTAAGTCTAATGTTCAAGATTTTCTACTGCTGGCAATCCCTCTGG 2565
 Qy 2427 TAAACAATGAAAAACATCTCTGAGCAATATTAATCTGCACTTTAGATAGAGT 2486
 Db 2566 TAAACATG-AAAAACATCTCTGAGTTATGTTAATCTGATCTTTAGAAATAGAA-- 2622
 Qy 2487 AACTTAATCTAGTCAATTTGAACCTGAATACATTTTCAATGATGAATAAGATAT- 2542
 Db 2623 ---ATAATGACCTCAGTCAAAAGTTCAATTAATTTTCAATTAATAAAGACATGAA 2679
 Qy 2543 -TATTTAAAGTATTCATGACATTTAATTA-TAAGTAGATTTTCAATATGCTT 2600
 Db 2680 CTATGTAATAAATAATTCATGACCAATATGTTATATTAACATGACTTCAATATTTACT 2739
 Qy 2601 AAGATTTATTCAGGAAACAAGTTCTCAATTTTACACAAAACTTTTACTATGATC 2660
 Db 2740 AAGAAATTAAGTCA-AAAAACAAGTTCTCAATTAATTAATGAATAATTTCTCAATGATC 2798
 Qy 2661 ACAGCTTTTCAATTAAGTCT-TTCTGATTAATCTGATTTTCAATTTATTAACAAGT 2719
 Db 2799 ATAAATCTTAATCTTTAATTAAGTCTTCTTAATTTTATTTTCTTAATTTCTGACAT 2858
 Qy 2720 AAATTAATTTAATTAACAATAAATTAATTTT-TACTTCAATGCTTACTTAATAGTA 2778
 Db 2859 AAAATTAATTTAATTAACAATAAATAAATTTTCTACTTCAATGCTTCAACATA 2918
 Qy 2779 TAAATCTTTTATTTTCTGAGGAAAGCAATCACTTTTATGATTAAGTAAAT 2838
 Db 2919 TAAATCTTTTATTTTCTGAGGAAAGTAAATTTTCT-ATAATTTAATAATTAAGTAAAT 2977
 Qy 2839 TAAGATTTATCTTAATTAATTAATTTTAAAGTTTCAATTAATTAAGATTAAGT 2898
 Db 2978 TAAATTTACTTAATTAATTAACAATAAAGTGTGAGGATTAAGTAAATTAAGTAC 3037
 Qy 2899 TAAGTATATGTTATTTTAAATTAAGTTTAAATTTTTCGATGCTCAATGACCTTT 2958
 Db 3038 ---AGTTATATTAATTCACAAATAGACTTTTGTCTTCAATGCTCAAAATATCTTTT 3093
 Qy 2959 ATTAATCTAATTAATTTCTTTTAAAGATTAAGTAAATTAAGTAAATTAAGTAAAT 3018
 Db 3094 ATTAATTTAATTAATTTCTTTTAAAGATTAAGTAAATTAAGTAAATTAAGTAAAT 3153
 Qy 3019 TAAACAATTAATCTTTTGTGAGCTTGTGACAGAGGTTTAACTCTGACAGAGAT 3078
 Db 3154 TAAACAATTAATCTGCTGAGCTTGTGAGAGAGAGGTTTAACTCTGACAGAGAT 3213
 Qy 3079 CATTTGAGCAAGAGCAATCAAGATCCGAGCACTTGTCTCATCAAAAGTTGAAT 3138
 Db 3214 CATTTGAGCAAGAGCAATCAAGATCCGAGCACTTGTCTCATCAAAAGTTGAAT 3273
 Qy 3139 ATAAAGGCACTTGAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAAT 3198
 Db 3274 ATAAAGGCACTTGAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAAT 3333
 Qy 3199 ACTGGGCAAGCAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAAT 3258
 Db 3334 ACTGACATGATTAATTTTGTGAGCTTAACTCAAAAGCAAAAGGATTAAGG 3393
 Qy 3259 AAGAGTAAAGCAAGGAAAGATTTGATTAATTTAAACC 3300

Db 3394 AAGAAACAGAAACAAGAAAAAGATTAATTAATTTAAATC 3435
 RESULT 2
 US-10-610-473-1
 ; Sequence 1, Application US/10610473
 ; Publication No. US20040030114A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Findlay, Robert
 ; TITLE OF INVENTION: MYOSTATIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AN
 ; TITLE OF INVENTION: METHODS FOR ITS USE
 ; FILE REFERENCE: PCI0448A
 ; CURRENT APPLICATION NUMBER: US/10/610, 473
 ; CURRENT FILING DATE: 2003-06-30
 ; PRIOR APPLICATION NUMBER: US/09/626, 959
 ; PRIOR FILING DATE: 2000-07-27
 ; PRIOR APPLICATION NUMBER: 60/146, 540
 ; PRIOR FILING DATE: 1999-07-30
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2482
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-10-610-473-1
 Query Match 15.5%; Score 525; DB 17; Length 2482;
 Best Local Similarity 61.5%; Pred. No. 2.8e-88;
 Matches 1256; Conservative 0; Mismatches 665; Indels 122; Gaps 21;
 Qy 1188 GATGAATTTAACATTAATCTCAACCCCTCGATGAAACAGACATTAAGTCAAGTA 1247
 Db 521 GAACACTTGAGATTAATTTGAAAGTAAAGGCTTGAACAAAGCAAGCCCTTCTGCT 580
 Qy 1248 CCACTTACGCTCTTAATTAAG-----ACAGCACTGATCAACAATTTACAT 1299
 Db 581 TCAAGTATTAATTTCTATGAAGGACCTACATTTAGCTTATATTTGCTAAATTAAT 640
 Qy 1300 GTATTAATTAATTAATGAGCCCAACCTAATGAGGAGCTGAAGGTTAAGAGTATC 1359
 Db 641 GCTCAACCCCTTTAGTTGAGAACTTAAGATTAAGAGAGCTTAAGTCTGCTCTT 700
 Qy 1360 TCATTTATTAATTAACAGATGAGCAATCTGAATCTGAATCTCAACTCCAGGA 1419
 Db 701 TGTATGACTTAAGAGGCAAAATTAAGATTTGAATCTGATTTTATTAATCTTCA 760
 Qy 1420 CTTGATCCCAATTTCTCAATTTCTAATCAAGTTAAAGGAAAGGATTTGATTT 1479
 Db 761 TCTGAGCTCAAAATGAGGAGTACAGCTTAAATAAATAAATCAAGATCA---ATT 816
 Qy 1480 GCTGAGAGTATAGGAGGATATGTTAC---AATTAATTAATTAAGATTAATNGT 1536
 Db 817 CTTGAGAGTATAGGAGGATATGTTATTAATTAATTAAGATTAAGATTAATNGT 876
 Qy 1537 TGAATAATTAATTTATCAACAATAAATTAATTTAAGCTGATTAATCTGCTCCGA 1596
 Db 877 AGAAATTAATTTTCCAGCAAAATCTGATTTAAGGCTGCTTAATATATGCTTGT 936
 Qy 1597 ACAAGACTATGAAATCTTCAAGGAGCTGTTGTTGTTCTTATCTTT 1656
 Db 937 --TAAAGAGGCGAGATCACTCAGGCTGCTGTTGTTGTTGTTCTTATCTTT 994
 Qy 1657 AATGAGGCAAACTTAATGATTAATTAAGGCA-----TTTTTCTCAAGATTAAG 1711
 Db 995 AATGAGGCAAACTTAATGATTAATTAAGGCACTTTTCTCAAGATTAAG 1054
 Qy 1712 ATACCTTTAAGATTTGATTAATTAATTAATTTTCAAGCTAAGTATGATTTA 1771
 Db 1055 ATGCTCTTAATAATTTGATTAATTAATTAATTTTCAAGCTAAGTATGATTTA 1114
 Qy 1772 GTGCACTGAGCAATTAATTT--AGTATCAATGTCGAAAGTATGATTAATTA 1828
 Db 1115 GTTCACTGAGGCAATTAATTTGATTAATTTGATTAATTTGATTAATTTA 1174

QY 1829 ATATTTGATATGAACCACTGACTCTCTGGGG-----AAAAAGTAAGATTAC 1880
 DB 1175 ATATTTATATCAAGCCACATAGTCTCTGGGGTGTGTGTAAGGAGATTAATTA 1234
 QY 1881 TCCTTGAAGAGTCTTGAAGTCTCCCAAAAGTAAGGAAAAATTAATCTCTGGGCTG 1940
 DB 1235 ATTTTAAGACTCCTGACGCTTCCCAACAGAGAGAGAA-----CTGTGGGCTG 1285
 QY 1941 GAAACAGCTTCTGTTCTTGTGCTGCTATATTTGTTAGGTTTAAATAGTTTCAATTGATT 2000
 DB 1286 GAAACGCTCTGTCCTCT--GCTGCTGTTGTTGCTAGCTCTTAAAGTTTCAACCCCAT 1343
 QY 2001 AGACCTTGTGGCTCCCAAGGTAAGTTGAGAGTTGATTCCTTCAAGGCGCACTTAAT 2060
 DB 1344 CGATCTTGTGGCTCTTAAGGCAAGGTAAGTTGATTCCTTCAAGGCGCACTTAAT 1403
 QY 2061 TTAGAGAACAAAAAGCCCATTTCTGCTCCAGACTTAACCCCAATCCCTGGCAGGTG 2120
 DB 1404 TCAGAGAACAAAAAGCACTTCTCTG-CTTAGCTTACCCAGATCCCTGGCAGGTG 1462
 QY 2121 TCTGCCCTCCGCTCAATGAGAACTGGCAAGAGTACTAGAGAGTGCACAGTACTA 2180
 DB 1463 TCTGCCCTCTGCTCAAAATGAGAGCGCTGGCAAGGGGTGCTAGCCTGTGACAGTAT---- 1518
 QY 2181 GAAAGTAGAAAAAGACTAGACACTAGTGAAGAGCAAGAAAAAGGCGCACTTCATGA 2240
 DB 1519 -----GGGAGCGCAACAAAGGACCCCTCTTCA 1547
 QY 2241 TGTGTTCTCTTCCCTTCTGTTGTCACAGTCTCCGATATTAATTACAGAGGTAGATA 2300
 DB 1548 ---TGGAGCTTCTTGTGTGTGTGCTAC--GGGAGCTGACATCATTCACAGAGAACAGCA 1603
 QY 2301 ACTACATTTTCTTTTACCACTGGAAGGCTGAGGAAACTTTTGA-----CCC 2351
 DB 1604 ---TTGCACTTACTGTCAGCCCTGGAAGTGAAGTCAAGAAATAGTCTCCAGGC 1660
 QY 2352 ATCAATAATTCACATCTCTTCACTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2411
 DB 1661 TACTTACAAAAATTCATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720
 QY 2412 ATATCTCTTGTGTAATAAACAATGAAAAACAATCTCTGCAATATTAATCTGCAAC 2471
 DB 1721 GGATTCGTGTGCTATCAAAAAAAGAAAAAAGAAAAAGAACACCTGCTTACTGCTGT 1780
 QY 2472 TTTAGAGTAGAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2531
 DB 1781 ATCTCTGAATAGAAAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1840
 QY 2532 ATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2586
 DB 1841 GTCAAGCCATTAAGCTTCAAAAGTAAGCAATGACAGAAAGCTGCCAGAGTGT 1900
 QY 2587 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2646
 DB 1901 TTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1960
 QY 2647 CTTTAACTAGATCAAGTCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2706
 DB 1961 TTTTAAAGGCTGACAACTGAGCTTTCTTAATTAATTAATTAATTAATTAATTAATTA 2020
 QY 2707 TTATACAAAGCTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2766
 DB 2021 TTTCT-----CTAATTAACCGACCTTAAGCAATTAATTTTCCCTCAATATTA 2068
 QY 2767 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2826
 DB 2069 GTTTTAAGTAACAAACAGACCTCAAGTCAAAAGATTAATTAATTTTGA-TTTTAAACA 2127
 QY 2827 TGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2886
 DB 2128 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2180

QY 2887 ATGAATTAAGATCTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2946
 DB 2181 ATTAATTAAGCTTAAGTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2239
 QY 2947 CATTAAGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3006
 DB 2240 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2299
 QY 3007 CACCTTGACTGTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3066
 DB 2300 TGCCCTGACTGTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2359
 QY 3067 GACAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3126
 DB 2360 GACAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2419
 QY 3127 CAAGTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3186
 DB 2420 CTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2479
 QY 3187 GTT 3189
 DB 2480 GTT 2482

RESULT 3

US-10-251-115-54
 ; Sequence 54, Application US/1025115
 ; Publication No. US2003012917A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Grobet, Luc; Georges, Michel; and Poncelet, Dominique
 ; TITLE OF INVENTION: MUTATIONS IN THE MYOSTATIN GENE CAUSING DOUBLE-MUSCLING IN MAMMALS
 ; FILE REFERENCE: 52836/00016
 ; CURRENT APPLICATION NUMBER: US/10/251,115
 ; CURRENT FILING DATE: 2002-09-20
 ; PRIOR APPLICATION NUMBER: 05/09/482,573
 ; PRIOR FILING DATE: 2000-01-13
 ; PRIOR APPLICATION NUMBER: 08/891,789
 ; PRIOR FILING DATE: 1997-07-14
 ; PRIOR APPLICATION NUMBER: 09/007,761
 ; PRIOR FILING DATE: 1998-01-15
 ; PRIOR APPLICATION NUMBER: PCT/IB98/01197
 ; PRIOR FILING DATE: 1998-07-14
 ; NUMBER OF SEQ ID NOS: 54
 ; SEQ ID NO 54
 ; LENGTH: 5790
 ; TYPE: DNA
 ; ORGANISM: bos taurus
 US-10-251-115-54

Query Match

Best Local Similarity 99.1%; Pred. No. 4,8e-85; Length 5790;

Matches 522; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 2864 TTTTAAAGTTTACATTAATAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2923
 DB 18 TATCAAAAGTTTACATTAATAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 77
 QY 2924 AGTTTAAATTTTGGCAGTGCACATTAAGCCTTATTAATTAATTAATTAATTAATTAAT 2983
 DB 78 AGTTTAAATTTTGGCAGTGCACATTAAGCCTTATTAATTAATTAATTAATTAATTAATTA 137
 QY 2984 AGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3043
 DB 138 AGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 197
 QY 3044 TGACAGACAGGTTTAACTTGTGACAGCAGATTCATTGTGAGCAAGAGCAATTCACA 3103
 DB 198 TGACAGACAGGTTTAACTTGTGACAGCAGATTCATTGTGAGCAAGAGCAATTCACA 257
 QY 3104 GATCCGACAGACCTTGTCTCATCAAGTTGGAATTAATAAAGCCACTTGGAAATCAGTA 3163
 DB 258 GATCCGACAGACCTTGTCTCATCAAGTTGGAATTAATAAAGCCACTTGGAAATCAGTA 317

QY 3164 TAAAGATTCAGTGTGGCAAGTGTCTCTGAGACTGGGCAAGCATTAAGTTGGCT 3223
Db 318 TAAAGATTCAGTGTGGCAAGTGTCTCT -AGACTGGGCAAGCATTAAGTTGGCT 376
QY 3224 TGGGCTTACTCAAAAGCAAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 3283
Db 377 TGGGCTTACTCAAAAGCAAAAGAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 436
QY 3284 GTATTGATTTTAAAGCAAGCAAAAGTCAATCTGTTTATTAATTAATTAATG 3343
Db 437 GTATTGATTTTAAAGCAAGCAAAAGTCAATCTGTTTATTAATTAATTAATG 496
QY 3344 TGATTTGTGTGCGCCAGTGGATCTGAATGAGAAAGAGAGAGAGAGAGAGAG 3390
Db 497 TGATTTGTGTGCGCCAGTGGATCTGAATGAGAAAGAGAGAGAGAGAGAGAG 543

RESULT 4
US-10-610-473-2

; Sequence 2, Application US/10610473
; Publication No. US20040030114A1
; GENERAL INFORMATION:
; APPLICANT: Findlay, Robert
; TITLE OF INVENTION: MYOSTATIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AN
; FILE REFERENCE: PCI0448A
; CURRENT APPLICATION NUMBER: US/10/610,473
; PRIOR APPLICATION NUMBER: US/09/626,959
; PRIOR FILING DATE: 2000-07-27
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Sus scrofa
US-10-610-473-2

Query Match 14.6%; Score 494.8; DB 17; Length 1537;
Best Local Similarity 83.7%; Pred. No. 1.1e-82;
Matches 676; Conservative 0; Mismatches 102; Indels 30; Gaps 9;

QY 2395 AGATCAATAGCTGCAATATCCTCTTGTAAATAAAGCAAAAGCAATCCTCGAG 2454
Db 747 AATTAAGAAAGCTGCAATATCCTCTTGTAAATAAAGCAAAAGCAATCCTCGAG 806
QY 2455 CAATATTATCTGCACTTTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2514
Db 807 CAATATTATCTGCACTTTAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 855
QY 2515 ATACAAATTTTCAATGATTAAGATTAATTAATTAATTAATTAATTAATTAATTA 2574
Db 856 GCAAGATTTTCACTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 915
QY 2575 TTA-AGTAGAGATTTTCAATGATTAAGATTAATTAATTAATTAATTAATTAATTA 2633
Db 916 GAATTAAGATTAATTTTCAATGATTAAGATTAATTAATTAATTAATTAATTAATTAAT 975
QY 2634 TATAGCAAGAAATCTTTTACTAGATCAAGCTTTTCAATTAAGTC-----TTCCGA 2687
Db 976 TATAGCTGAATATTTTACTAGATCAAGCTTTTCAATTAATTAATTAATTAATTAATTA 1035
QY 2688 ATAAATCTGATTTT-TCTAATTAACAGAGCTAAAGATTAATTAATTAATTAATTAAT 2746
Db 1036 TTAATCTGATTTTCTGATTAACAGAGCTAAAGATTAATTAATTAATTAATTAATTAAT 1095
QY 2747 TATTTTACTTCAATGCTTAAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2806
Db 1096 TCTTTTCTCAATGCTTAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1155

QY 2807 --GCATTAACATTTTAAAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2864
Db 1156 AGACATTTCACTTTTAAAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1211
QY 2865 TTTTAAAGTTTCAATTAAGATTAAGATTAAGT---GATATGTTATTTGTTAT 2921
Db 1212 TTTTAAAGTTTCACTTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1271
QY 2922 AAAGTTTAAATTTTGGCATGTCAGATACAGCTTTTATTAATTAATTAATTAATTTCTTT 2981
Db 1272 ATAGATTTTAAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTT 1331
QY 2982 TAAAGATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3041
Db 1332 TATAGATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1391
QY 3042 TGTACAGACAGAGGTTTAACTCTGACAGAGAGATTAATTAATTAATTAATTAATTAAT 3101
Db 1392 TGTACAGACAGAGGTTTAACTCTGACAGAGAGATTAATTAATTAATTAATTAATTAAT 1451
QY 3102 CAGATCCGACAGACATTTGTCATCAAAAGTTGAGATTAATTAATTAATTAATTAATTAAT 3161
Db 1452 TAGATCCGACAGACATTTGTCATCAAAAGTTGAGATTAATTAATTAATTAATTAATTA 1509
QY 3162 TATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3189
Db 1510 TATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1537

RESULT 5
US-10-610-473-2/c

; Sequence 2, Application US/10610473
; Publication No. US20040030114A1
; GENERAL INFORMATION:
; APPLICANT: Findlay, Robert
; TITLE OF INVENTION: MYOSTATIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AN
; FILE REFERENCE: PCI0448A
; CURRENT APPLICATION NUMBER: US/10/610,473
; PRIOR APPLICATION NUMBER: US/09/626,959
; PRIOR FILING DATE: 2000-07-27
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Sus scrofa
US-10-610-473-2

Query Match 9.8%; Score 332.8; DB 17; Length 1537;
Best Local Similarity 63.9%; Pred. No. 2.6e-52;
Matches 555; Conservative 0; Mismatches 302; Indels 11; Gaps 3;

QY 2485 GTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2544
Db 1223 GAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1164
QY 2545 TTTAAAGTAATTCATGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2604
Db 1163 AATAGCTTTTCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1104
QY 2605 ATTATTCAGGAGAAAGAGTTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2664
Db 1103 AAAAGAAATTTTATTTGCTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1044
QY 2665 TCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2724
Db 1043 AGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 988
QY 2725 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2784

RESULT 7
 US-10-027-632-120231
 ; Sequence 120231, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; PRIORITY FILING DATE: 2002-04-30
 ; PRIORITY APPLICATION NUMBER: US 60/218,006
 ; PRIORITY FILING DATE: 2000-07-12
 ; PRIORITY APPLICATION NUMBER: US 60/198,676
 ; PRIORITY FILING DATE: 2000-04-20
 ; PRIORITY APPLICATION NUMBER: US 60/193,483
 ; PRIORITY FILING DATE: 2000-03-29
 ; PRIORITY APPLICATION NUMBER: US 60/185,218
 ; PRIORITY FILING DATE: 2000-02-24
 ; PRIORITY APPLICATION NUMBER: US 60/167,363
 ; PRIORITY FILING DATE: 1999-11-23
 ; PRIORITY APPLICATION NUMBER: US 60/156,358
 ; PRIORITY FILING DATE: 1999-09-28
 ; PRIORITY APPLICATION NUMBER: US 60/146,002
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 120231
 ; LENGTH: 942
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-027-632-120231

Query Match 6.6%; Score 224.8; DB 17; Length 942;
 Best Local Similarity 63.5%; Pred. No. 3.9e-32;
 Matches 589; Conservative 1; Mismatches 258; Indels 80; Gaps 13;
 402 TTTCATGAGGAAAAATTTTCTCTGAGTCCCTGCGACAGCCCTCTGCTTAAATTTG 461
 14 TTTTATGAGGAACTTTTCCCTAGGCAATACAGCCCTT-----TTAAATTTG 68
 462 ACAAATATATACATACATGAAAGTACTATTTAATCTGATTTTAACTGATGAGTT 521
 69 TAAATATATACACAACTTATGCA-----TCTTAACATTTTATGATGATGCG 119
 522 TGGATTAAGTACCTTACCTTCTGCGAATGATGCACTTCACTTCACTTCACTT 581
 120 TGGATTAAGTACCTTCACTTCTGCGAATGATGCACTTCACTTCACTTCACTT 179
 582 TTGATCTTCTGAGTGAATGAAATCTGCACTTATGCAAGTATGCTTCCCATCAGC 641
 180 TTGATCTTCTGAGTGAATGAAATCTGCACTTATGCAAGTATGCTTCCCATCAGC 641
 642 GCACCTCTCTGCAAGCCCGCGGACCAACATCTTCTTTCATGACAGTATCTGTGC 701
 232 --GTGCTCAACCCAGTCCCTGCGAACAACATGCTTCACTTCACTTCACTTCTGTGA 289
 702 CTAGGAACA-----CAGCCCTTGAACATGCTTGTGGGGGCACTGT---A 744
 290 TTAGAAACATAGAACACACAGCAGACGTCAGACATATGCTTGGGGGCACTTCTTAAA 349
 745 AGCAACAGATCACTCCCTACCGCACCAAAATGCACAAAATATATAAAGATGGG 804
 350 AGCAAAATCATATAGAGAGACATATAAAGAAAGACAAATATGTAAGAACTGGAC 409
 805 CATATGATGCAAAAGAGGCTGTGCT--AAGTATGAGGGCTGAAC--AAGCAGAG 859
 410 TAAATATAGCGGAAAGAGGAAATTTGTTTATATGATGAGACTGAACAAAGGCGAA 469
 860 AATTGATAGGTTGATCTGAGTGGATCTGTGTGTGG--AAGCTCAAAATTTTTCAT 917
 470 CCTGCTGTTGATCCTCATCTGAGAACCTGTGCTGATGATGAGATCAAAATTTTTCG 529

Qy 918 TGTTCGTGATACGCAAAATGCTTATTAAGACATGTAAGATTAAGATTAA 977
 Db CGCTGTGTAATGCGAGGAATGACTGAAAAA-----TGCTAGGATATATGATTCAATG 585
 Qy 978 GATTAATCTCAGCAAGACATATAATGTGACAGCGGATCCATGATTAACAGACTGAC 1037
 Db 586 ATAAATTTTACCAAGTGGCAAAATTTGTAATACGAATATATTAAGAAATGTC 645
 Qy 1038 CAT--GTGGAATGATATATCTTGTTCCTTATTCAGCAGCAGTGAAGAAAGCGC--- 1092
 Db 646 TGTTAGCGGAATATACAACTTTGTTTCACTTTTTCAGGATGATTAAGAAAGCTGG 705
 Qy 1093 -----TCAAGGCTGCTTACACCACTTTTACTAGAGAGTACCTATGCTAGT 1141
 Db 706 AACATGCTTGTATACCTGCTGCTTGTGAAATTTTACTTCTCAGAGGCTGAAAGT 765
 Qy 1142 CGG-----TAGCTGGCAATTAACAACTGAAGCACTTCTAGTTCAATGAGAGA 1189
 Db 766 AGGCTTTGTGACTGATATATGCGAGCTATGACCTGAAGCACTTCTAGTTCAATGAGAGA 825
 Qy 1190 TGAATTTAACCATATATCTCAACCCCTGCGATGAAACAGAGACTAAGTACTCAAGTACC 1249
 Db 826 TAAATTTAAGCATATATCTCAACCCCTTCTGATTAACAAAGAGCAAGCACTCAATGCC 885
 Qy 1250 AGTTATCAGTCACTTACTATATGACAG 1277
 Db 886 AGTTATCATTACTTACTATATGACAG 913

RESULT 8
 US-10-366-345-27
 ; Sequence 27, Application US/10366345
 ; Publication No. US20030224501A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, et al.
 ; TITLE OF INVENTION: Bone Morphogenetic Protein Polynucleotides, Polypeptides and
 ; FILE REFERENCE: PT189
 ; CURRENT APPLICATION NUMBER: US/10/366,345
 ; PRIORITY FILING DATE: 2003-02-14
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 27
 ; LENGTH: 2823
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-366-345-27

Query Match 5.6%; Score 190; DB 17; Length 2823;
 Best Local Similarity 91.0%; Pred. No. 2.1e-25;
 Matches 202; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
 3168 AGATTCACTGCTGCGAAGTTGCTCTCAAGCTGGGCGAGCATTAAGTTGCTGGC 3227
 Db 1 AGATTCACTGCTGCGAAGTTGCTCTCAAGCTGGGCGAGCATTAAGTTGCTGGC 60
 Qy 3228 GTTATCTCAAAAGCAAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGTAAAGTAA 3287
 Db 61 ATTACTCAAAAGCAAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGTAAAGTAA 120
 Qy 3288 TGATTTTAAACATGCAAAATCTGTTATATTAATTAATTAATTAATTAATTAAT 3347
 Db 121 TGATTTTAAATCATGCAAAATCTGTTATATTAATTAATTAATTAATTAATTAAT 180
 Qy 3348 TGTGCTGCGCCAGTGAATCTGAATGAAACAGGAGACAGA 3389
 Db 181 TGTGCTGCTCAGTGAATCTGAATGAAACAGGAGACAGA 222

RESULT 9
 US-10-610-473-3
 ; Sequence 3, Application US/10610473
 ; Publication No. US2004003011A1

```

; GENERAL INFORMATION:
; APPLICANT: Findlay, Robert
; TITLE OF INVENTION: MYOSTATIN REGULATORY REGION, NUCLEOTIDE SEQUENCE DETERMINATION AND
; TITLE OF INVENTION: METHODS FOR ITS USE
; FILE REFERENCE: PC10448A
; CURRENT APPLICATION NUMBER: US/10/610,473
; CURRENT FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US/09/626,959
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/146,540
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-610-473-3

```

```

Query Match      5.1%; Score 174.4; DB 17; Length 200;
Best Local Similarity 92.0%; Pred. No. 5.8e-23;
Matches 184; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2390 AGTCAATGATCAGCTCACCCTTGACTGTACACAAATACCTGTTGGTGTGACAG 3049
    |||||
DB 1 AGTCAATGATCAGCTTGCCTGACTGTACACAAATGCTGCTGGTGTGACAG 60

QY 3050 ACAGGGTTTAACTCTGACACCGAGATTCTATTGGAGCAGACCCATACAGATCCC 3109
    |||||
DB 61 ACAGGGTTTAACTCTGACACCGAGATTCTATTGGAGCAGACCCATACAGATCCC 120

QY 3110 GACGACACTTCTTCATCAAGTTGGAATATTAAGACCACTGGAATACAGTATTAAG 3169
    |||||
DB 121 GACGACACTTCTTCATCAAGTTGGAATATTAAGACCACTGGAATACAGTATTAAG 180

QY 3170 ATTCAGCTGTGTGGCAGTT 3189
    |||||
DB 181 ATTCAGCTGTGTGGCAGTT 200

```

```

RESULT 10
US-10-251-115-1
; Sequence 1, Application US/10251115
; Publication No. US20030129171A1
; GENERAL INFORMATION:
; APPLICANT: Grobet, Luc; Georges, Michel; and Poncelet, Dominique
; TITLE OF INVENTION: MUTATIONS IN THE MYOSTATIN GENE CAUSING DOUBLE-MUSCLING IN MAMMAL
; FILE REFERENCE: 52836/00016
; CURRENT APPLICATION NUMBER: US/10/251,115
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/482,573
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 08/891,789
; PRIOR FILING DATE: 1997-07-14
; PRIOR APPLICATION NUMBER: 09/007,761
; PRIOR FILING DATE: 1998-01-15
; PRIOR APPLICATION NUMBER: PCT/IB98/01197
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 54
; SEQ ID NO 1
; LENGTH: 1196
; TYPE: DNA
; ORGANISM: Bos taurus
; US-10-251-115-1

```

```

Query Match      4.0%; Score 134.4; DB 15; Length 1196;
Best Local Similarity 99.3%; Pred. No. 4e-15;
Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3255 AGGAAGAAGTAAGACAGGAGAAAGATTGTATTTTAAACCATGCAAAACTGCA 3314
    |||||
DB 1 AGGAAGAAGTAAGACAGGAGAAAGATTGTATTTTAAACCATGCAAAACTGCA 60

```

```

QY 3315 AATCTGTTTATTAATTAACCTATTATGCTGATTTGCTGGCCCACTGATCTGAATGA 3374
    |||||
DB 61 AATCTGTTTATTAATTAACCTATTATGCTGATTTGCTGGCCCACTGATCTGAATGA 120

QY 3375 GAACAGGAGCAGAG 3390
    |||||
DB 121 GAACAGGAGCAGAG 136

```

```

RESULT 11
US-10-251-115-3
; Sequence 3, Application US/10251115
; Publication No. US20030129171A1
; GENERAL INFORMATION:
; APPLICANT: Grobet, Luc; Georges, Michel; and Poncelet, Dominique
; TITLE OF INVENTION: MUTATIONS IN THE MYOSTATIN GENE CAUSING DOUBLE-MUSCLING IN MAMMAL
; FILE REFERENCE: 52836/00016
; CURRENT APPLICATION NUMBER: US/10/251,115
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/482,573
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 08/891,789
; PRIOR FILING DATE: 1997-07-14
; PRIOR APPLICATION NUMBER: 09/007,761
; PRIOR FILING DATE: 1998-01-15
; PRIOR APPLICATION NUMBER: PCT/IB98/01197
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 54
; SEQ ID NO 3
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Bos taurus
; US-10-251-115-3

```

```

Query Match      3.9%; Score 132.8; DB 15; Length 1240;
Best Local Similarity 98.5%; Pred. No. 8e-15;
Matches 134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3255 AGGAAGAAGTAAGACAGGAGAAAGATTGTATTTTAAACCATGCAAAACTGCA 3314
    |||||
DB 1 AGGAAGAAGTAAGACAGGAGAAAGATTGTATTTTAAACCATGCAAAACTGCA 60

QY 3315 AATCTGTTTATTAATTAACCTATTATGCTGATTTGCTGGCCCACTGATCTGAATGA 3374
    |||||
DB 61 AATCTGTTTATTAATTAACCTATTATGCTGATTTGCTGGCCCACTGATCTGAATGA 120

QY 3375 GAACAGGAGCAGAG 3390
    |||||
DB 121 GAACAGGAGCAGAG 136

```

```

RESULT 12
US-09-854-867-233
; Sequence 233, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL, H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 233
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(653)
; OTHER INFORMATION: mer82

```

```

; NAME/KEY: misc feature
; LOCATION: (186)..(187)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (544)..(544)
; OTHER INFORMATION: n is a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (557)..(557)
; OTHER INFORMATION: n is a, c, g or t
; NAME/KEY: misc feature
; LOCATION: (567)..(567)
; OTHER INFORMATION: n is a, c, g or t
US-09-854-867-233
```

Query Match 3.7%; Score 126.2; DB 10; Length 653;

Best Local Similarity 70.9%; Pred. No. 1.1e-13;

Matches 222; Conservative 1; Mismatches 81; Indels 9; Gaps 4;

```

Qy 119 AGTGACTCTGATTTATTCAGAGTGTTATGTTTACCAATCACTGCCAACA-TGAATPA 177
Db 2 AGTTGATCTCTATATTCACAGTAGTTATGTTTAAAGTCAACGCGACGACTGAATTA 61
Qy 178 GTGATTCCTGAACACAGTCCCTAGGGAGACTACAGAGTTAGTTCCGTCAGCCTCTGG 237
Db 62 GCGAATCTGAACACAGTCTCTAGAGGAATACAGATTAGTTCCGTCAGCCTCTGG 121
Qy 238 TC---ACGTTTGTAAACAATCAATTAATTAATCTGTTTGTGTCATTTCTGTTTAA 294
Db 122 TCACACATTTTCAATCAACCATCATATATATATCTGTTTATGTTGTTCTGTTTAA 181
Qy 295 AAATATCTTTTATTAATAGTACTGTAATCTTCAATTTGTTTCAACCAAAAGGC 354
Db 182 AGACNMC-TTATTTAATATATATATGTTGATTATTAACCTGAACTCAGCCAAAC-- 238
Qy 355 CCTATTACTGGAAGCCCTGGAATGAAGCTTACATTAACACATTTGTTTCTATGAGAA 414
Db 239 --CACTAATCACTGATGCTGATGAGCTTATCTAACAACATATTTTCTCATTAAGTA 296
Qy 415 AAATTTTCTTC 427
Db 297 CATCAGACCTTC 309
```

RESULT 13

US-10-786-970A-233

; Sequence 233, Application US/10786970A

; Publication No. US2005006449A1

; GENERAL INFORMATION:

; APPLICANT: JOAN, KNOUL

; APPLICANT: ROGAN, PETER

; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING

; FILE REFERENCE: 30307

; CURRENT APPLICATION NUMBER: US/10/786,970A

; CURRENT FILING DATE: 2004-02-24

; PRIOR APPLICATION NUMBER: US/09/573,080

; PRIOR FILING DATE: 2000-05-16

; NUMBER OF SEQ ID NOS: 479

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 233

; LENGTH: 653

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: repeat_region

; LOCATION: (1)..(653)

; OTHER INFORMATION: mer82

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: n is a, c, g or t

; PUBLICATION INFORMATION:

```

; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J, Malchiewicz, J, Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
US-10-786-970A-233
```

Query Match 3.7%; Score 126.2; DB 19; Length 653;

Best Local Similarity 70.9%; Pred. No. 1.1e-13;

Matches 222; Conservative 1; Mismatches 81; Indels 9; Gaps 4;

```

Qy 119 AGTGACTCTGATTTATTCAGAGTGTTATGTTTACCAATCACTGCCAACA-TGAATPA 177
Db 2 AGTTGATCTCTATATTCACAGTAGTTATGTTTAAAGTCAACGCGACGACTGAATTA 61
Qy 178 GTGATTCCTGAACACAGTCCCTAGGGAGACTACAGAGTTAGTTCCGTCAGCCTCTGG 237
Db 62 GCGAATCTGAACACAGTCTCTAGAGGAATACAGATTAGTTCCGTCAGCCTCTGG 121
Qy 238 TC---ACGTTTGTAAACAATCAATTAATTAATCTGTTTGTGTCATTTCTGTTTAA 294
Db 122 TCACACATTTTCAATCAACCATCATATATATATCTGTTTATGTTGTTCTGTTTAA 181
Qy 295 AAATATCTTTTATTAATAGTACTGTAATCTTCAATTTGTTTCAACATTTGTTTCAACAAAGGC 354
Db 182 AGACNMC-TTATTTAATATATATGTTGATTATTAACCTGAACTCAGCCAAAC-- 238
Qy 355 CCTATTACTGGAAGCCCTGGAATGAAGCTTACATTAACACATTTGTTTCTATGAGAA 414
Db 239 --CACTAATCACTGATGCTGATGAGCTTATCTAACAACATATTTTCTCATTAAGTA 296
Qy 415 AAATTTTCTTC 427
Db 297 CATCAGACCTTC 309
```

RESULT 14

US-09-859-211-13

; Sequence 13, Application US/09859211

; Patent No. US20020157125A1

; GENERAL INFORMATION:

; APPLICANT: Lee, Se-Jin

; APPLICANT: McPherson, Alexandra C.

; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8

; FILE REFERENCE: 07265/144001

; CURRENT APPLICATION NUMBER: US/09/859,211

; CURRENT FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: 09/019,070

; PRIOR FILING DATE: 1998-02-05

; PRIOR APPLICATION NUMBER: 08/862,445

; PRIOR FILING DATE: 1997-05-23

; PRIOR APPLICATION NUMBER: 08/847,910

; PRIOR FILING DATE: 1997-04-28

; PRIOR APPLICATION NUMBER: 08/795,071

; PRIOR FILING DATE: 1997-02-05

; PRIOR APPLICATION NUMBER: 08/525,596

; PRIOR FILING DATE: 1995-10-26

; PRIOR APPLICATION NUMBER: PCT/US94/03019

; PRIOR FILING DATE: 1994-03-18

; PRIOR APPLICATION NUMBER: 08/033,923

; PRIOR FILING DATE: 1993-03-19

; NUMBER OF SEQ ID NOS: 51

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 2743

; TYPE: DNA

; ORGANISM: Homo sapiens

QY 1607 TCGAATCTCTGAGGACATCTGTTGTGTCGTTTCTTAACTTTAATGATGGG 1666
DB 405 GCACAGATCACCCTCGGGGTGCTGCTGTGTGTCTGTTTCTTCACTTTAATGCTGGC 346
QY 1667 AAATCTAATGATATGATGAAGGCA-----TTTTTCTCAAGAGATGATGATCTTTA 1721
DB 345 AAATCTAATGATATGATGAAGGCACTTTTCTTCTCAAGATATGATGCTCTT 286
QY 1722 AGAATTTGATGAATAATGATTAATCTTTTCAGGCTACTGATGATTTGATGACAG 1781
DB 285 AAATTTGATGAATAATGATTAATCTTTTCAGGCTACTGATGATTTGATGACAG 226
QY 1782 GCGATAATT---AGTGTACAAATGTGGAAAGTAGTACCTAAATAATATTTGATA 1838
DB 225 GCAATTAATTTGGGTGTACTGTACAGAAATGTGTGACCTTAATAATATTTGATA 166
QY 1839 TGAACCACTGCACTCTCTGGGG-----AAAAAGTAATGATTAATCTCTTTAGA 1890
DB 165 CAACCCACATAGTCTCTGGGGGTGTGTGTAAAGGAGTATGATTAATTAATCTTAAGA 106
QY 1891 GTCTTAGCTTCCCAAAAGTAGTAGAAATAATCTCTGTGGCTTGAAACAGCTT 1950
DB 105 CTCTCAGCTTCCCAACAGAGAGGAGAA-----CTCTGTGGCTTGAAAGCTCT 55
QY 1951 CTGTTCTGTGCTGATATTTGTTTAGTTTAAATAGTTCAATTTAGTAGACT 2006
DB 54 CTGTCCCT--GCTGCTGTGTTGTTCAGCTTTTAAGATTCAACCCATTCGATCT 1

RESULT 2
LOCUS BX971880/c 829 bp DNA linear GSS 05-JUL-2004
DEFINITION Forward strand read from insert in 5'HPT insertion targeting and chromosome engineering clone MHPN10117, genomic survey sequence.
ACCESSION BX971880
VERSION BX971880.1 GI:49703303
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
FEATURES
source location/Qualifiers
1..829
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN10117"
/clone_lib="MHPN"

ORIGIN
Query Match 6.9%; Score 235.2; DB 9; Length 829;
Best Local Similarity 68.9%; Pred.No.5.4e-37;
Matches 452; Conservative 0; Mismatches 168; Indels 36; Gaps 8;

QY 1370 TTACACAGTGGCAACTGATCTGAACCTCAGCTCTATCCACTCCAGAGCTGAGATCC 1429
DB 639 TTAGAGAGGCAAAATTTGAGATTGAACCTCAGCTTTATTGACTCTTCAGCTCAGCTCA 580
QY 1430 CAATTGCTACACAAATCTTAATCAAGTTAAAGGAAAGATTGATTGCTCAGAGT 1489
DB 579 CAATGCACTACAGCTTAATAAAAAAATCAACAGATCA-----ATTCTCTGAGGT 524
QY 1490 GTATAGGGGCAATATGTTAC---AATTATACATTACAAAGATTATATGTTGAAAAATTA 1546
DB 523 AATATGACAGCATGTGTATGATTAATATGACATCGAAAGAAATTTATGACAGAAAAATGA 464

QY 1547 AATTATCAACAATAATACTTTATAGCCTGATCTAATCTGCTCCGCAAGACATA 1606
DB 463 AATTTCAGACAAATCTGATTTATAGGCTGCTCTAATATGTCCTGTA--TAAAGAG 406
QY 1607 TCGAATCTCTGAGGACATCTGTTGTGTCGTTTCTTAACTTTAATGATGGG 1666
DB 405 GCACAGATCACCCTCGGGGTGCTGCTGTGTGTCTGTTTCTTCACTTTAATGCTGGC 346
QY 1667 AAATCTAATGATATGATGAAGGCA-----TTTTTCTCAAGAGATGATGATCTTTA 1721
DB 345 AAATCTAATGATATGATGAAGGCACTTTTCTTCTCAAGATATGATGCTCTT 286
QY 1722 AGAATTTGATGAATAATGATTAATCTTTTCAGGCTACTGATGATTTGATGACAG 1781
DB 285 AAATTTGATGAATAATGATTAATCTTTTCAGGCTACTGATGATTTGATGACAG 226
QY 1782 GCGATAATT---AGTGTACAAATGTGGAAAGTAGTACCTAAATAATATTTGATA 1838
DB 225 GCAATTAATTTGGGTGTACTGTACAGAAATGTGTGACCTTAATAATATTTGATA 166
QY 1839 TGAACCACTGCACTCTCTGGGG-----AAAAAGTAATGATTAATCTCTTTAGA 1890
DB 165 CAACCCACATAGTCTCTGGGGGTGTGTGTAAAGGAGTATGATTAATTAATCTTAAGA 106
QY 1891 GTCTTAGCTTCCCAAAAGTAGTAGAAATAATCTCTGTGGCTTGAAACAGCTT 1950
DB 105 CTCTCAGCTTCCCAACAGAGAGGAGAA-----CTCTGTGGCTTGAAAGCTCT 55
QY 1951 CTGTTCTGTGCTGATATTTGTTTAGTTTAAATAGTTCAATTTAGTAGACT 2006
DB 54 CTGTCCCT--GCTGCTGTGTTGTTCAGCTTTTAAGATTCAACCCATTCGATCT 1

RESULT 3
LOCUS CR242965/c 835 bp DNA linear GSS 06-JUL-2004
DEFINITION Forward strand read from insert in 5'HPT insertion targeting and chromosome engineering clone MHPN35104, genomic survey sequence.
ACCESSION CR242965
VERSION CR242965.1 GI:50021814
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
FEATURES
source location/Qualifiers
1..835
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN35104"
/clone_lib="MHPN"

ORIGIN
Query Match 6.8%; Score 231.8; DB 9; Length 835;
Best Local Similarity 69.2%; Pred.No.2.7e-36;
Matches 445; Conservative 0; Mismatches 162; Indels 36; Gaps 8;

QY 1370 TTACACAGTGGCAACTGATCTGAACCTCAGCTCTATCCACTCCAGAGCTGAGATCC 1429
DB 629 TTAGAGAGGCAAAATTTGAGATTGAACCTCAGCTTTATTGACTCTTCAGCTCAGCTCA 570
QY 1430 CAATTGCTACACAAATCTTAATCAAGTTAAAGGAAAGATTGATTGCTCAGAGT 1489
DB 569 CAATGCACTACAGCTTAATAAAAAAATCAACAGATCA-----ATTCTCTGAGGT 514

```

QY 1490 GTATGAGGCGATATGTAC--AATTATACATTAACAAGATTATATGTTGAAAAATTA 1546
DB 513 ATATAGCAGCATGTGTATATGATTAATTAATGACATCAAAAAGAAATCTATGACAAAAATGA 454
QY 1547 ATTTATCAACAAATTAATACTTATAGCTGATCTATATCTGCTCCGCAACAAAGACTA 1606
DB 453 ATTTTCAGACAAATCTGATCTTATAGGCTCTTAATATGTTCTTTGA--TAAAGAG 396
QY 1607 TCTGAATCTTTCAGGCGATCTGTTGTCGTGTTTCTTAACTTAAATGANGGAC 1666
DB 395 GCGAGATACCTCCAGGGTGTCTGCTTGTGTCTGTTTCTTCAATCTTAAATGATGCTCTTA 336
QY 1667 AAATCTAATGATTAATGTAAGGCCATTTT-----CTCAAGATGATGATACCTCTTA 1721
DB 335 AAATCTAATGATTAATGTAAGGCCATTTT-----CTCAAGATGATGATACCTCTTA 276
QY 1722 AGAATTAATGATAATGATTAATCTTTCAGGCTACTGATGTTGATTTAGTGCATGAG 1781
DB 275 AAAATTAATGATAATGATTAATCTTTCAGGCTACTGATGTTGATTTAGTGCATGAG 216
QY 1782 GCAATTAATGATAATGATTAATCTTTCAGGCTACTGATGTTGATTTAGTGCATGAG 1838
DB 215 GCAATTAATGATAATGATTAATCTTTCAGGCTACTGATGTTGATTTAGTGCATGAG 156
QY 1839 TGAACCACTGACCTCTCTTGGGG-----AAAAAGTAATGATTAATCTCTTAGAGA 1890
DB 155 CAAGCACAATGATGCTCTTGGGGGTGTGTAAAGGGAATGATTAATCTCTTAGAGA 96
QY 1891 GTCTTACTGATCCCAAAAGTAATGATAATTAATCTCTGTCGCTCGAAGACGTT 1950
DB 95 CTCTCAGCTTCCCAAAAGTAATGATAATTAATCTCTGTCGCTCGAAGACGTT 45
QY 1951 CTGTTCTTGTGCTGCTATATTTGTTAGTTTAAATGTTCA 1993
DB 44 CTGTCCTT--GCTGCTGTTGTTTTCAGCTCTTAAAGTTCA 4

RESULT 4
LOCUS CR198021 615 bp DNA linear GSS 06-JUN-2004
DEFINITION Forward strand read from insert in 5'Hprt insertion targeting and
Chromosome engineering clone MHPN400j17, genomic survey sequence.
ACCESSION CR198021.1 GI:4976870
VERSION GSS: genome survey sequence; MICER.
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weijden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
source
1. 615
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN400j17"
/clone_lib="MHPN"
ORIGIN
Query Match 6.4%; Score 218.6; DB 9; Length 615;
Best Local Similarity 69.7%; Pred. No. 1.2e-33;
Matches 387; Conservative 0; Mismatches 143; Indels 25; Gaps 6;

```

```

QY 1442 AATCTTAATCAAGTTAAAGGAAAGAAAGATTGATTTGCTCAGAGTATAGGGGANT 1501
DB 489 AGT-----CTAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 434
QY 1502 ATGTAC--AATTATACATTAACAAGATTATATGTTGAAAAATTAATTAATTAACA 1558
DB 433 GTTAAATATATTAATGATGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 374
QY 1559 AATAAATCTTAATTAAGCTGATTAATCTGCTCGCAACAAAGCATCTGAATCTT 1618
DB 373 AATCTGATTAATTAAGCTGCTGCTTAATTAATGTTCTTTGA--TAAAGAGGCCAGATCCT 316
QY 1619 CAGGCAATCTGTTGTTGTCGTTTCTTAACTTAAATGATGAGCAAAATCTAATGCA 1678
DB 315 CAGGCTGCTCTTGTGTCTGTTTCTTCACTTAAATGATGAGCAAAATCTAATGCA 256
QY 1679 TTAATTAAGGCA-----TTTTTCTCAAGATGATGATACCTTTAAGAAATTTGATGA 1733
DB 255 TTAATTAAGGCACTTTTCTTCTCAAGATGATGATACCTTTAAGAAATTTGATGA 196
QY 1734 AAATGCAATTAATTTTCAAGCTACTGATGTTGATTTGATGATGATGATGATGATGATGAT 1791
DB 195 AAATGCAATTAATTTTCAAGCTACTGATGTTGATTTGATGATGATGATGATGATGATGAT 136
QY 1792 -AGTGTACAATGTCGAAAGTAGTACCTAAATAATTAATTTGATGATGATGATGATGATGATGAT 1850
DB 135 GTGTACTGTACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 76
QY 1851 CTCTCTTGGGG-----AAAAAGTAATGATTAATCTCTTAAAGATCTTAAAGCTTTC 1902
DB 75 GTCTCTTGGGGGTGTGTAAAGGAGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 16
QY 1903 CCCAAATTAATGAG 1917
DB 15 CCAAAAGAGAGAG 1

RESULT 5
LOCUS AG325355 680 bp DNA linear GSS 02-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-113C02.T7, genomic survey
sequence.
ACCESSION AG325355
VERSION AG325355.1 GI:4789346
KEYWORDS
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-cho,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@sc.riken.jp, URL:http://hgp-gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBACe3.6

```


sciences center (GSC), Yokohama Institute

[illegible]

Location/Qualifiers
1. .681

/organism="Mus musculus"
 /mol_type="rRNA"
 /db_xref="taxon:10090"
 /clone="D330014D01"
 /tissue_type="heart"
 /dev_stage="13 days embryo"
 /lab_host="DH10B"
 /clone_lib="Riken full-length enriched, 13 days embryo
 heart"
 /note="Site 1: Sali; Site 2: Banhi; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN, Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5].
 GAGGAGAGAGCGGCCGCACTGAGTTTCTTTTCTTTTNN 3', cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length b
 cap-trapper. Second strand cDNA was prepared with the
 primer adaptor of sequence [5].
 GAGGAGAGAGATTCGAGTAATTAATTAATCCCCCCCCCC 3'. cDNA
 was cleaved with BanHI and XhoI. Vector: a modified
 Bluescript KS(+) after bulk excision from lambda FLC I.

61 CATTACTCAAAAGCAAAAGAGAAATAAGACCAAGGCAAAAAAAAAAAGATTGTCGTGAT - 119

TTAAACCATGCAAAACTGCAATCTGTATTATATTACCTATTATGCTGA 334

Db 258 AAAGCACAA 250

LOCUS	1101 bp	DNA	linear	GSS 04-JUN-1999
CNS000EVL				
Drosophila melanogaster genome survey sequence T7 end of BAC:				
DEFINITION				
CNS000EVL/c				
RESULT 10				

ORIGIN

Query Match	3.2%	Score 107.8;	DB 9,	Length 1101;
Best Local Similarity	35.4%	Pred. No. 3.7e-11;		
Matches 224;	Conservative 126;	Mismatches 278;	Indels 5;	Gaps 2;

[illegible]

QY	2607	TTATTCAGGGGAAACAGTTCTCAAAATTAATGCGAAAATCTTTTACATGATCAACAGTC	2666
Db	797	WTATTTATATATATATAMMTTMAITMTTMAAATAATATATMTATTTATTTATTTTWT	738
QY	2667	TTTTCAATTAAGTCTCTCGAATAAATCTGTATTTTCTAAATTATACAAGCTAAATAAT	2726
Db	737	TTATTTTATTAATTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTAAATTA	678
QY	2727	TTTAAATATACAATAA--AATTAATTTTACTCAAAATGCTTACTTAATATGATTAATAAT	2784
Db	677	TTTAAATAMWMAATWMTATTTTATATATTAATTAATTTTATTTTATTTTATTTTAAAAA	618
QY	2785	CATTTTATTTTTCAGGGGAAAAGATATCACTTTTAAAGTATGAAGTGTATTAATTAAGAT	2844
Db	617	TATTTTTTATATTAATTAATWMTATTTATATMAATATATMTTWTAAATTTTAAATATWT	558
QY	2845	TTATTCACCTTAATTAATTTTAAAGTTTCACATATAAGATGAATTAAGATCTAAGTG	2904
Db	557	WTTTAAATATATTTTTTTTTTAAATMAAMWTTMAAAATAATWMAAMWMAAAAAATTTTWT	498
QY	2905	TATATGTTATGTTAATAAAGTTTATATTTTTT	2937
Db	497	TTMAATTAATAATTTTWTTTTMAAATTTTWWKKKKK	465

RESULT	11
B95505/c	
LOCUS	
DEFINITION	571 bp DNA linear GSS 26-JUN-1998 CIT-HSP-216SE24.TF CIT-HSP Homo sapiens genomic clone 216SE24, genomic survey sequence.
ACCESSION	B95505
VERSION	B95505.1 GI:2977842
KEYWORDS	GSS.
SOURCE	Homo sapiens (human)

ORIGIN	
Query Match	3.2%; Score 107.4; DB 8; Length 571;
Best Local Similarity	70.0%; Pred. No. 4.3e-11;

QY 109 GATTGATAGAGCTGACTCTCATTTATTCACAGTGGTTATGTTTCAACCAATCATCTGCCAA 168
 DB 251 GTTGAATAACAGCTGATTCATTTTCAAGTAGTCTTCTGTAAAGTCAACAGGAA 310
 QY 169 C-ATGAAATAGATTTCTGTAACCACTGCCCCAGGGAATCAAGATTTAGTCCGT 227
 DB 311 CGCTGAATTAACAAATTTGTAACCTTCTCCCAAAATAATGAGGGTGAATTTCTGCG 370
 QY 228 CAGCCTCTGTGAC---GTTTTGTAAACCAATCATTAATTAACCTTGTGTTGATGCA 283
 DB 371 AGGCTCTGTGACCAACATTTTGTGCATCATTAATTAATTTGTTATGTT 430
 QY 284 TTTCTGTTTAAATATCTTATTTAATACGTAATCTTAATTTCAACATTTGGTTCAC 343
 DB 431 TTTCTGTTTAAAGACA-CCTTATGTAATACATCTGTTGATTCATTAACATTTGAACCTAC 489
 QY 344 AACCAAAAGCCCTTATTAACGTAAGCCCTGATATGAGCTTAACATTAACACATTTGTTT 403
 DB 490 AGCCAAATG---CACTGTAACTCATCTGTAATGACATATTAACACACATTTTTC 545
 QY 404 TCTATGAGAAATTTTCTTCTGATC 431
 DB 546 TCTGGAGGACATCATCAGCTTTCTTTC 573

RESULT 14

AQ120350 487 bp DNA linear GSS 22-SEP-1998
 LOCUS HS_3010_B2_B01_T7 CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate=3010 Col=2 Row=D, genomic survey
 sequence.

ACCESSION AQ120350
 VERSION AQ120350
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 487)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 3010 row: D column: 2
 Class: BAC ends
 High quality sequence stop: 487.

FEATURES

SOURCE

1. 487

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=3010 Col=2 Row=D"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
 E-Coli DH10B"

ORIGIN

Query Match 3.1%; Score 104; DB 8; Length 487;
 Best Local Similarity 61.5%; Pred. No. 2.1e-10;

Matches 201; Conservative 0; Mismatches 121; Indels 5; Gaps 2;
 QY 121 CTGACTCTATTAATTCACAGTGGTTATGTTTCAACCAATCATCTGCCAAATGAAGT 180
 DB 16 CTGGACATTCATTTTCTCGATGATTAATTCACAAAGTTGCCAAACACTGATAGCT 75
 QY 181 ATTCCTGAACACATGCCCCATAGGGAACCTAACAGGTTAGATTCCTGAGCTGTGCA 240
 DB 76 TATACTGAACCATGATCTTATGGGAATTAACAGGGTATGCTCTTGAAGCTTATGTC 135
 QY 241 CGTTTGTGTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
 DB 136 CATCTGTGCAACAAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 194
 QY 301 CTTTATTAATACGTAATCTTAATCTTCAACATTTGTTTCAACAAAGGCTTAT 360
 DB 195 CTTTATTAATTAATGATGTTGTTGATTCATTAACATTTGAACCTGACCTAAGG---CACT 250
 QY 361 AACTGAAGCCCTGAATGAACCTTAACATTAACACACATTTGTTTCTATGAGGAAATTT 420
 DB 251 AGAGGATGATGATTCATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 310
 QY 421 TTCCTGATCTCTGCAACGCTTCTT 447
 DB 311 TGTGATTAAGAAATGCAATAGTGTCTT 337

RESULT 15

AG124176 624 bp DNA linear GSS 04-NOV-2001
 LOCUS AG124176
 DEFINITION Pan troglodytes DNA, clone: PTB-134A20.R, genomic survey sequence.

ACCESSION AG124176
 VERSION AG124176.1 GI:1665341
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 1
 AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
 Torok,Y., Watanabe,H. and Sakaki,Y.
 BAC end sequences of library PTB
 Unpublished
 2 (bases 1 to 624)

TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Aaso Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suenho-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.
 PRIMERS
 Sequencing: M13Rev

LIBRARY Vector : PKS145
 R.Site 1 : SacI
 R.Site 2 : SacI

FEATURES

SOURCE

1. 624

/organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-134A20.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 3.1%; Score 103.4; DB 9; Length 624;

Best Local Similarity 66.5%; Pred. No. 2.8e-10;
Matches 163; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

```
QY 66 ATTTCTTGAGTACCTCTTAATTCATTCATGCTGACCTCGGAGATTGATAGCTGAC 125
    |||||
Db 361 ATCTGTTGTCATACCGCATCTTTCACCTCCAGATATTTTATTAATAGACTAT 302
    |||||
QY 126 TCTCATATCAAGTGTATGTTCTACCAATCACTGCCACATG-AATTAAGTATTC 184
    |||||
Db 301 TCTTTTATGACAGTATGTTATGTTCTATTAATTCACCTGCACCAACCAATTGTGAATA 242
    |||||
QY 185 CTGAACCACTGCCCTAGGGGAGTACAGGTAGATTCCGTCAGCCTCTGTACGTT 244
    |||||
Db 241 CTGAACATATCACTCTAGGGGAAACATAGGTGAGTTCTGCAAGGCTGTATACATT 182
    |||||
QY 245 TTGTTAACCAATCAATTAATTAACCTTGTGTTGTGATTTCTGTTTAAATATCTTT 304
    |||||
Db 181 TTTATCAACCAATCAATTAATTAACCTTGTGTTGTGATTTCTGTTTAAAGACACCTTA 122
    |||||
QY 305 ATTTA 309
    |||
Db 121 TTAA 117
```

Search completed: March 31, 2005, 14:38:27
Job time : 6904 secs